

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 15:25:26 / Search time 40.3032 Seconds
(without alignment)
2847.400 Million cell updates/sec

Title: US-10-009-945-2
Perfect score: 3884
Sequence: 1 GGSISIKRLTVLCARLANK.....EKLYEKLTAVETGPAVE 723

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

- A_Geneseq_19Jun03:*
- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
 - 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
 - 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
 - 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
 - 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
 - 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
 - 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
 - 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
 - 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
 - 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
 - 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
 - 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
 - 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
 - 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
 - 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
 - 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
 - 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
 - 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
 - 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
 - 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
 - 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
 - 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
 - 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
 - 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3884	100.0	723	AA1980.DAT:*
2	3878	99.8	722	AA1981.DAT:*
3	3869	99.6	722	AA1982.DAT:*
4	3801	77.3	748	AA1983.DAT:*
5	3001	77.3	748	AA1984.DAT:*
6	2996	77.1	804	AA1985.DAT:*
7	2236	57.6	514	AA1986.DAT:*
8	2069	53.3	1035	AA1987.DAT:*
9	1614.5	41.6	335	AA1988.DAT:*

10	1461	37.6	766	AA1989.DAT:*
11	1436.5	37.0	832	AA1990.DAT:*
12	1345	34.6	869	AA1991.DAT:*
13	1318	33.9	854	AA1992.DAT:*
14	1301	33.5	852	AA1993.DAT:*
15	1259	32.4	927	AA1994.DAT:*
16	1259	32.4	927	AA1995.DAT:*
17	1245.5	32.1	739	AA1996.DAT:*
18	1245.5	32.1	739	AA1997.DAT:*
19	1245.5	32.1	739	AA1998.DAT:*
20	1245.5	32.1	739	AA1999.DAT:*
21	1245	32.1	739	AA2000.DAT:*
22	1239	31.9	922	AA2001.DAT:*
23	1239	31.9	922	AA2002.DAT:*
24	1239	31.9	949	AA2003.DAT:*
25	1239	31.9	949	AA2004.DAT:*
26	1238.5	31.6	923	AA2005.DAT:*
27	1228.5	31.6	995	AA2006.DAT:*
28	1227	31.6	724	AA2007.DAT:*
29	1226.5	31.6	759	AA2008.DAT:*
30	1225	31.5	870	AA2009.DAT:*
31	1225	31.5	870	AA2010.DAT:*
32	1223.5	31.5	834	AA2011.DAT:*
33	1223	31.5	854	AA2012.DAT:*
34	1212	31.2	474	AA2013.DAT:*
35	1212	31.2	506	AA2014.DAT:*
36	1193.5	30.7	733	AA2015.DAT:*
37	1157.5	29.8	1082	AA2016.DAT:*
38	1151	29.6	1071	AA2017.DAT:*
39	1150.5	29.6	818	AA2018.DAT:*
40	1047	27.0	683	AA2019.DAT:*
41	1047	27.0	683	AA2020.DAT:*
42	1047	27.0	684	AA2021.DAT:*
43	1023	26.1	380	AA2022.DAT:*
44	1012	26.1	375	AA2023.DAT:*
45	1002	25.8	1562	AA2024.DAT:*

ALIGNMENTS

RESULT 1	AA1980.DAT:*
ID	AA1980.DAT:*
AC	AA1980.DAT:*
DT	20-APR-2001 (first entry)
XX	
DE	Amino acid sequence of a human Smurf1 polypeptide.
XX	
KW	Smurf1; Smurf2; Sma2 signal transduction; bone morphogenetic protein; BMP;
KW	transforming growth factor-beta; human; TGF-beta; Chondrogenesis;
KW	osteogenesis; blood differentiation; cartilage formation; hair growth;
KW	neural tube patterning; retinal development; heart induction;
KW	morphogenesis; tooth formation; gamete formation.
XX	
OS	Homo sapiens.
XX	
PN	WO200077168-A2.
XX	
PD	21-DEC-2000.
XX	
PF	12-JUN-2000; 2000WO-US16250.
XX	
PR	11-JUN-1999; 99US-0138969.
XX	
PA	(UNY) UNIV NEW YORK STATE RES FOUND.
XX	
PA	(HSCR-) HSC RES & DEV LP.
XX	
PI	Thomsen GH, Wzana J;
XX	
DR	WPI; 2001-071267/08.

Human protein ubiq
Candida albicans e
Aspergillus fumiga
Mortine E3 ubiquiti
Human E3 ubiquitin
Human KIAAN ligase
Human NBD4 long f
Novel human diagno
Atrophin-1 interac
Human atrophin-1 i
Atrophin-1 interac
Human ZG6B1 prote
Human homolog of D
Human signal trans
Drosophila melanog
Drosophila melanog
Novel human diagno
KIAA439 protein.
Novel human gene,
Aspergillus fumiga
Human ubiquitin pr
Nedd-4-like ubiquit
Diseased kidney ti
Human cell growth,
Amino acid sequenc
Novel human protei
Human clone 811a
Drosophila melanog
Novel human diagno
Human protein ubiq
Novel human protei
Human ubiquitin pr
Nedd-4-like ubiquit
Human WPI HECT E3
Human WPI HECT E3
Novel human diagno

DR N-PSDB; AAF24852.

XX Novel isolated Smurf protein useful for inhibiting bone morphogenic

PT protein or tumor growth factor-beta activation pathway, for treating

PT cancer and to block osteogenesis, hair growth, tooth formation

XX

XX Claim 6; Fig 10; 107pp; English.

XX

CC The present sequence represents a human Smurf1 polypeptide. The

CC specification also describes a Smurf2 polypeptide. Smurf polypeptides

CC are negative regulators of Smad signal transduction, and antagonists of

CC bone morphogenic protein (BMP) or transforming growth factor-beta

CC (TGF-beta) signalling pathway. Expression of Smurf1 in a cell is useful

CC for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf

CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood

CC differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth

CC formation, gamete formation and a wide variety of tissue and organ

CC formation processes, and hinder the regeneration, growth, maintenance,

CC etc., of bone and other tissues that are dependent on the BMP pathway.

CC The polypeptide is useful for screening for various drugs and/or

CC antibodies that can either enhance the BMP pathway, or inhibit it.

XX

SQ Sequence 723 AA;

Query Match 100.0%; Score 3884; DB 22; Length 723;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSIKIRLTLVCAKRLAKKDFRLLPDPFAKIVDVGSGQCHSTDTVKNTLDPKNNOHYDL 60

DB 1 GSSSIKIRLTLVCAKRLAKKDFRLLPDPFAKIVDVGSGQCHSTDTVKNTLDPKNNOHYDL 60

QY 61 YVAGTDSITISVNMHKKIKKQAGAGLGCVRLLSNAISLKDQYQRLDLCKLNPSTDA 120

DB 61 YVAGTDSITISVNMHKKIKKQAGAGLGCVRLLSNAISLKDQYQRLDLCKLNPSTDA 120

QY 121 VRGQIVVSLQTRIRITGSGVVDRCGLLENSTGYVEDSPGRPLSCFMEEPAPLTDSTGA 180

DB 121 VRGQIVVSLQTRIRITGSGVVDRCGLLENSTGYVEDSPGRPLSCFMEEPAPLTDSTGA 180

QY 181 AAGGNCRFVESPQDQRLQAQRLRNPDVRSGLQTPQNRPHGQSPELPEGYEORTVOG 240

DB 181 AAGGNCRFVESPQDQRLQAQRLRNPDVRSGLQTPQNRPHGQSPELPEGYEORTVOG 240

QY 241 QVYFLHTQVGSVWHPRIPLDINSVNCDELGPLPPGMEVRSVTSGRILYFVDHNNRTTQF 300

DB 241 QVYFLHTQVGSVWHPRIPLDINSVNCDELGPLPPGMEVRSVTSGRILYFVDHNNRTTQF 300

QY 301 TDPRLHHIMHQCQLKEPQPLPLPESGSLDEBELPAQRYERDLVQKLKVRHELSLQCP 360

DB 301 TDPRLHHIMHQCQLKEPQPLPLPESGSLDEBELPAQRYERDLVQKLKVRHELSLQCP 360

QY 361 QAGHCRIVSRBEIFESRQIMKMRPKDLKRLMKVFRGEGGLDYGVAEMLYLCH 420

DB 361 QAGHCRIVSRBEIFESRQIMKMRPKDLKRLMKVFRGEGGLDYGVAEMLYLCH 420

QY 421 MNPYYGLFOYSTDNLYMQLINPDSSINPDHLSYFVFRINGLAVFGHYNGFTYF 480

DB 421 MNPYYGLFOYSTDNLYMQLINPDSSINPDHLSYFVFRINGLAVFGHYNGFTYF 480

QY 481 YKQLGKPIQLSLESVDPELAKSLVLIENDITPYLDHTFCVENNAGRLIOHELKXNG 540

DB 481 YKQLGKPIQLSLESVDPELAKSLVLIENDITPYLDHTFCVENNAGRLIOHELKXNG 540

QY 541 RNVPVTEENKKEYRVLVVMRFRGIEAOFLALQKFNFLIQHLLKPFDDKELELLIGG 600

DB 541 RNVPVTEENKKEYRVLVVMRFRGIEAOFLALQKFNFLIQHLLKPFDDKELELLIGG 600

QY 601 LDKIDLNDKSNTRLKCAVADSNIVMFWQAVETPEERARLLQFTVGTSTVPLQGFKA 660

DB 601 LDKIDLNDKSNTRLKCAVADSNIVMFWQAVETPEERARLLQFTVGTSTVPLQGFKA 660

QY 661 LOGSTGAAGPRLFTIHLIDANTDLPKATCFNRIDIPRESYEKLYEKLTAVEETCGF 720

DB 661 LOGSTGAAGPRLFTIHLIDANTDLPKATCFNRIDIPRESYEKLYEKLTAVEETCGF 720

QY 721 AVE 723

DB 721 AVE 723

RESULT 2

AAE32724

ID AAE32724 standard; Protein, 722 AA.

XX

XX AAE32724;

AC

XX

DT 24-MAR-2003 (first entry)

XX

DE E3 ubiquitin ligase SMURF1 protein.

XX

XX Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;

KW polioyellitis; HIV; measles; protein therapy; E3 ubiquitin ligase;

XX enzyme.

XX

OS Unidentified.

XX

PN WO200290549-A2.

PD

PD 14-NOV-2002.

XX

PF 12-MAR-2002; 2002MO-IB02106.

XX

PR 12-MAR-2001; 2001US-275224P.

PR 31-JUL-2001; 2001US-308258P.

PR 07-DEC-2001; 2001US-340170P.

XX

PA (PROT-) PROTEOLOGICS LTD.

XX

PI Greener T, Moskowitz H, Reiss Y, Alroy I;

PI WPI; 2003-111976/10.

DR

DR N-PSDB; AAD50458.

XX

PT New protein complex comprising HECT-RCC1, viral maturation scaffolding

PT protein (VMSF), and/or HIV gag protein, useful for treating viral

PT infections, such as lymphosarcoma, HIV, hepatitis, polioyellitis, measles,

PT or Ebola

XX

PS Disclosure; Fig 14; 150pp; English.

XX

CC The invention relates to a method for modulation of viral maturation.

CC The invention also provides an isolated protein complex comprising a

CC HECT-RCC1 selected from HERT-NW, HECT-RCC1, Gag protein, Gag late

CC domain, p13, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2, STAM2B,

CC VHS-UIM, GTPase, E2 enzyme, teg101, cullin, HRC1, HRC2, HRC3, Nedd4

CC -like protein or clathrin. The complexes, proteins, antibodies and

CC methods are useful for treating viral infections, such as lymphosarcoma,

CC human immunodeficiency virus (HIV), hepatitis, polioyellitis, measles,

CC or Ebola and for inhibiting budding in a subject. They are also useful

CC in diagnostic assays for determining whether a cell is infected with a

CC virus and for characterizing the nature, progression and/or infectivity

CC of the infection. The invention is also useful in protein therapy. The

CC present sequence is E3 ubiquitin ligase SMURF1 protein used to illustrate

CC the method of the invention.

XX

SQ Sequence 722 AA;

Query Match 99.8%; Score 3878; DB 24; Length 722;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSSSIKIRLTLVCAKRLAKKDFRLLPDPFAKIVDVGSGQCHSTDTVKNTLDPKNNOHYDL 61

DB 1 GSSSIKIRLTLVCAKRLAKKDFRLLPDPFAKIVDVGSGQCHSTDTVKNTLDPKNNOHYDL 60

```

QY 62 VGTKDSTITISVNMHKKIHKKQAGFLGCVRLLSNAISRDKDTGYORLDCIKNPSDTDAV 121
DB 61 VGTKDSTITISVNMHKKIHKKQAGFLGCVRLLSNAISRDKDTGYORLDCIKNPSDTDAV 120
QY 122 RGOIVVSLQTRDRIGTGGSVVDGRGLLENEGTVYEDSGGRPLSGFMEEPAYDTSTGAA 181
DB 121 RGOIVVSLQTRDRIGTGGSVVDGRGLLENEGTVYEDSGGRPLSGFMEEPAYDTSTGAA 180
QY 182 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRHGHQSPLEBEGYQRTTVGQ 241
DB 181 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRHGHQSPLEBEGYQRTTVGQ 240
QY 242 VYFLHTQTVSTWHDPRIPRDLSVNCDELGRLPPMEVRSVSGRIYFVDNNNTTQFT 301
DB 241 VYFLHTQTVSTWHDPRIPRDLSVNCDELGRLPPMEVRSVSGRIYFVDNNNTTQFT 300
QY 302 DPLRHHIMNHQCCQKESQPLPSPSGSLEDEBLPAQRYERDLVOKLKVLRHELSLQCPQ 361
DB 301 DPLRHHIMNHQCCQKESQPLPSPSGSLEDEBLPAQRYERDLVOKLKVLRHELSLQCPQ 360
QY 362 AGHCRIVSREBEIFEEBYSYROIIMKQRPKDLKKRLMVKFRGSEGLDYGVARWLYLLCHEM 421
DB 361 AGHCRIVSREBEIFEEBYSYROIIMKQRPKDLKKRLMVKFRGSEGLDYGVARWLYLLCHEM 420
QY 422 LNPYYGLFQYSTDNIVYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPEFY 481
DB 421 LNPYYGLFQYSTDNIVYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPEFY 480
QY 482 KQLGKPIQLSDLESVDPPELHKSIVMLLENDIPVLDHIFCVENHAFGRILQHELKPKGR 541
DB 481 KQLGKPIQLSDLESVDPPELHKSIVMLLENDIPVLDHIFCVENHAFGRILQHELKPKGR 540
QY 542 NVPTBENKKEVYVLYNWMRFMRGIEAQLALQGFNEILPQHLNRPDQKELEIIGGL 601
DB 541 NVPTBENKKEVYVLYNWMRFMRGIEAQLALQGFNEILPQHLNRPDQKELEIIGGL 600
QY 602 DKIDNPKSNTRLAKHCVADSNIVRMFWQAVETPDEERRARLLQFTGSTRVPLQGFRL 661
DB 601 DKIDNPKSNTRLAKHCVADSNIVRMFWQAVETPDEERRARLLQFTGSTRVPLQGFRL 660
QY 662 QGSTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLEYEKLITAVEETGFA 721
DB 661 QGSTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLEYEKLITAVEETGFA 720
QY 722 VE 723
DB 721 VE 722

```

```

PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSQ INC.
PI Tang YT, Liu C, Dirmancic RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
DR N-PSDB; AAK52994.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 362-363; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC
CC Note: Records for SEQ ID NO 2110 (AAK52591), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 722 AA;

```

```

RESULT 3
AAK79861
ID AAK79861 standard; Protein; 722 AA.
XX
AC AAK79861;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3507.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX

```

```

QY 2 GSSIKIRLTVLCAKNLAKKOPFLRPPFAKIYVDSGGCHSTDTYKNTLDPKMNQHYDLY 61
DB 1 GSSIKIRLTVLCAKNLAKKOPFLRPPFAKIYVDSGGCHSTDTYKNTLDPKMNQHYDLY 60
QY 62 VGTKDSTITISVNMHKKIHKKQAGFLGCVRLLSNAISRDKDTGYORLDCIKNPSDTDAV 121
DB 61 VGTKDSTITISVNMHKKIHKKQAGFLGCVRLLSNAISRDKDTGYORLDCIKNPSDTDAV 120
QY 122 RGOIVVSLQTRDRIGTGGSVVDGRGLLENEGTVYEDSGGRPLSGFMEEPAYDTSTGAA 181
DB 121 RGOIVVSLQTRDRIGTGGSVVDGRGLLENEGTVYEDSGGRPLSGFMEEPAYDTSTGAA 180
QY 182 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRHGHQSPLEBEGYQRTTVGQ 241
DB 181 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRHGHQSPLEBEGYQRTTVGQ 240
QY 242 VYFLHTQTVSTWHDPRIPRDLSVNCDELGRLPPMEVRSVSGRIYFVDNNNTTQFT 301
DB 241 VYFLHTQTVSTWHDPRIPRDLSVNCDELGRLPPMEVRSVSGRIYFVDNNNTTQFT 300
QY 302 DPLRHHIMNHQCCQKESQPLPSPSGSLEDEBLPAQRYERDLVOKLKVLRHELSLQCPQ 361
DB 301 DPLRHHIMNHQCCQKESQPLPSPSGSLEDEBLPAQRYERDLVOKLKVLRHELSLQCPQ 360
QY 362 AGHCRIVSREBEIFEEBYSYROIIMKQRPKDLKKRLMVKFRGSEGLDYGVARWLYLLCHEM 421
DB 361 AGHCRIVSREBEIFEEBYSYROIIMKQRPKDLKKRLMVKFRGSEGLDYGVARWLYLLCHEM 420
QY 422 LNPYYGLFQYSTDNIVYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPEFY 481

```

```

Db 421 LNPYGLFQYSTEINIMQLINPDSSINPDHLSYFFHVGRIKGLAFVGHYINGGFTVPFY 480
Qy 482 KQLGKPIQLSDPESDPELHKSLVWILENDITPLDHTFCVHNAPFRILOHCLKNGR 541
Db 481 KQLGKPIQLSDPESDPELHKSLVWILENDITPLDHTFCVHNAPFRILOHCLKNGR 540
Qy 542 NVPTEENKKEYRLLVYVNRFPNRGIEAGFLALQKGFNELIPQHLKPPDQKELELIIGL 601
Db 541 NVPTEENKKEYRLLVYVNRFPNRGIEAGFLALQKGFNELIPQHLKPPDQKELELIIGL 600
Qy 602 DKTLNDMKSNTFLKCVADSNIVKFMQAVETPDEERRARLLQFTVMSSTVPVQGFAL 661
Db 601 DKTLNDMKSNTFLKCVADSNIVKFMQAVETPDEERRARLLQFTVMSSTVPVQGFAL 660
Qy 662 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPYSEYKLYEKLTAVEETCGFA 721
Db 661 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPYSEYKLYEKLTAVEETCGFA 720
Qy 722 VE 723
Db 721 VE 722

```

RESULT 4

AAW13384 ID AAW13384 standard; Protein; 748 AA.

AC AAW13384;

DT 10-JUL-1997 (first entry)

DE Human protein ubiquitin ligase publi.

KM Protein ubiquitin ligase; publi; cdc25 phosphatase; CDK kinase; p53; cell cycle; transgenic animal.

OS Homo sapiens.

PN MO9712962-A1.

PD 10-APR-1997.

PF 04-OCT-1996; 96MO-US15930.

PR 04-OCT-1995; 95US-0539205.

PA (COLD-) COLD SPRING HARBOR LAB.

PI Beach D, Caligiuri M, Nefsky B;

DR WPI; 1997-226206/20.

DR N1-SDB; AAT47040.

PT Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved in cdc25 phosphatase and p53 ubiquitination, and regulate cell growth and proliferation

PS Claim 1; Page 74-77; 108pp; English.

XX Human protein ubiquitin ligases publi (AAW13384), pub2 (AAW13385) and CC pub3 (AAW13386) are homologues of fission yeast publi (AAW13387) and CC were identified from cDNA clones (AAW17040-42) obtcd. e.g. from a CC keratinocyte cDNA library. Pub polypeptides can be produced in CC transfected host cells. They can control the steady state level of CC cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2) CC dephosphorylation and the steady state level of p53 (controlling CC the degree of cell cycle regulation of p53). They can regulate CC cell or tissue differentiation, or cell growth or proliferation by CC affecting other proteins, can be a specific (ant)agonist of wild-type CC type protein function and may be used as immunogens to elicit a CC specific immune residue.

Seq Sequence 748 AA;

Query Match 77.3%; Score 3001; DB 18; Length 748;

Best Local Similarity 74.6%; Pred. No. 6, 4e-286;

Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

```

Qy 5 IKIRLTVLCANLAKKOPFRLDPDPAKIVVDSGGCHSTDTVKNLDPKMNQHYDLYVK 64
Db 12 VILRLTVLCANLAKKOPFRLDPDPFAVVDGSGCHSTDTVKNLDPKMNQHYDLYVK 71
Qy 65 TDSITISVWNNKKIHKKQAGAGLGCVALLSNAISRLKDTGYRLDCKLNPSPTDAVQ 124
Db 72 SDSVTISVWNNKKIHKKQAGAGLGCVALLSNAISRLKDTGYRLDCKLNPSPTDAVQ 131
Qy 125 IVVSLQTRDRIIGTGGVVDGRLLENE-----GTVY----- 155
Db 132 IVVSLQSRDRIIGTGGVVDGRLLENE-----GTVY----- 191
Qy 156 --EDSGRPRPLSCMEERFAPYTDSTGAAGGNCRFVESPQDQRLAQRHPDVRGSL 213
Db 192 ASERISPPRPLSCFVDENTPISTNGATCG-----OSDPRLEARRVRSQRHNTV 242
Qy 214 QTPQRRPHGOSPELPEGEYEQRTVQGVFLHTQGVSTWHDPRIPRDLNSVNCDELGP 273
Db 243 ---SRTHLTPPDLPEGEYEQRTVQGVFLHTQGVSTWHDPRVPRDLNSINCELG 298
Qy 274 LPPGWEVASTVSGRIYFDHNKRTTQTDPR---LHINNHQQLKEPSQPLPESBGS 329
Db 299 LPPGWEINRTAGRVFVDHNNRTQTDPRLSANLHVNRQNLKDDQDQGV---S 354
Qy 330 L---EDEELPAQRYERDLYVQLKRLHLSLQOPQACHRIEVSREIEPESYRQIMKOR 386
Db 355 LCPDDELTVPKRYKRDLYVQLKRLHLSLQOPQACHRIEVSREIEPESYRQIMKOR 414
Qy 387 PDKLKKRLMKVFRGEGIDYGVARREYLILCHEMLNPYGLQYSTDTNLYMLQINPDS 446
Db 415 PDKLKKRLMKVFRGEGIDYGVARREYLILCHEMLNPYGLQYSTDTNLYMLQINPDS 474
Qy 447 INPDLSYFFHVGRIKGLAFVGHYINGGFTVPFYKQLGKPIQLSDPESDPELHKSLV 506
Db 475 VNPEHLSTFHFVGRIMGAAPVGHYIDGGFTLPPYKQLGKSLTIDDMELVDDLSNLV 534
Qy 507 WLENDITPVLDHTFCVHNAPFRILOHCLKNGRNVVTEENKKEYRLLVYVNRFRGI 566
Db 535 WLENDITGVLDHTFCVHNAPFRILOHCLKNGKSIPIVNEENKKEYRLLVYVNRFRGI 594
Qy 567 EAQFLALQKGFNELIPQHLKPEPOCKLELIIGSLDKIDNDMKSNTRLKHCYVADSNIV 626
Db 595 EAQFLALQKGFNEVLPQHLKTFEKELELIIGSLGKIDNDMKSNTRLKHCYVADSNIV 654
Qy 627 WFMQAVETPDEERRARLLQFTVGSSTRVPLQGFALQGSTGAAGPRLFTIHLIDANTDNL 686
Db 655 WFMQAVETPDEERRARLLQFTVGSSTRVPLQGFALQ---GAAGPRLFTIHLIDANTDNL 711
Qy 687 KATCFNRIDIPYSEYKLYEKLTAVEETCGFAVE 723
Db 712 KATCFNRIDIPYSEYKLYEKLTAVEETCGFAVE 748

```

RESULT 5

AAW13477 ID AAW13477 standard; Protein; 748 AA.

AC AAW13477;

DT 20-APR-2001 (first entry)

DE Amino acid sequence of a human Smurf2 polypeptide.

DB Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;

KM transforming growth factor-beta; human; TGF-beta; chondrogenesis;

KW osteogenesis; blood differentiation; cartilage formation; hair growth;

KW neural tube patterning; retinal development; heart induction;

KM morphogenesis; tooth formation; gamete formation.
 XX Homo sapiens.
 OS
 XX MO200077168-A2.
 PN
 XX 21-DEC-2000.
 PD
 XX 12-JUN-2000; 2000WO-US16250.
 PF
 XX 11-JUN-1999; 99US-0138969.
 PR
 XX (UNYX) UNIV NEW YORK STATE RES FOUND.
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Thomsen GH, Wrana J;
 XX
 DR MPI, 2001-071267/08.
 DR N-PSDB; AAF24853.
 XX
 XX Novel isolated Smurf protein useful for inhibiting bone morphogenic
 PT protein or tumor growth factor-beta activation pathway, for treating
 PT cancer and to block osteogenesis, hair growth, tooth formation -
 XX
 PS Claim 10; Fig 12; 107bp; English.
 XX
 CC The present sequence represents a human Smurf2 polypeptide. The
 CC specification also describes a Smurf1 polypeptide. Smurf polypeptides
 CC are negative regulators of Smad signal transduction, and antagonists of
 CC bone morphogenic protein (BMP) or transforming growth factor-beta
 CC (TGF-beta) signalling pathway. Expression of Smurf1 in a cell is useful
 CC for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf
 CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood
 CC differentiation, cartilage formation, neural tube patterning, retinal
 CC development, heart induction and morphogenesis, hair growth, tooth
 CC formation, gamete formation and a wide variety of tissue and organ
 CC formation processes, and hinder the regeneration, growth, maintenance,
 CC etc., of bone and other tissues that are dependent on the BMP pathway.
 CC The polypeptide is useful for screening for various drugs and/or
 CC antibodies that can either enhance the BMP pathway, or inhibit it.
 CC
 XX Sequence 748 AA;
 XX
 Query Match 77.3%; Score 3001; DB 22; Length 748;
 Best Local Similarity 74.6%; Pred No. 6.4e-286;
 Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;
 QY 5 IKRLITVLCANLAKKDFRLPDPFAKIVVDSGGCHSTDTVKTLDPKMNQHYDLYGK 64
 DB 12 VKRLITVLCANLAKKDFRLPDPFAKIVVDSGGCHSTDTVKTLDPKMNQHYDLYGK 71
 QY 65 TDSITISVWNHKKIHKKQAGFLGCVRLISNAISRLKDTGVRDLCLNPSDDAVRGQ 124
 DB 72 SDSVTISVWNHKKIHKKQAGFLGCVRLISNAISRLKDTGVRDLCLNPSDDAVRGQ 131
 QY 125 IVVSLQTRDRIGTSGVVDRCGLLENE-----GVVY----- 155
 DB 132 IVVSLQTRDRIGTSGVVDRCGLLENE-----GVVY----- 191
 QY 156 --EDSGRPLSLFMEZEPAYDSTGAAGGNCVFESPSDQQLQARLNDVGRSL 213
 DB 192 ASEYSSPGRPLSEFVDENTPISGTNGATG-----QSSDPRLLARRRSQRHRYM 242
 QY 214 QTPQNRPHGSHSELEGEQRTTVQGVYFLHTGTGVSFWHPRIIPRLDSVNCDELGP 273
 DB 243 ---SRTHLTPDPLEGEQRTTVQGVYFLHTGTGVSFWHPRIIPRLDSVNCDELGP 298
 QY 274 LPPGWEVRSVSGRIIFYVDNNRTTOFTDPR---LHIANHQQCKLPPSGPLPSEGS 329
 DB 299 LPPGWEIRNTATRVYFVDNNRTTOFTDPRLSANLHLVLRNQDLKQDQOQVY---S 354
 QY 330 L--EDELPAQYEDLVQKLVLRHELSLOPQAGHCRLEVSREELFEESYQIMKR 386

DB 355 LCPDTECLTVPRYKRDLYQKILNRQELSQQCPQAGHCRLEVSREELFEESYQIMKR 414
 QY 387 PKDLKRLMVFRRGEELDYGVAREMLYLLGCEMNPYGYLFQYSTDNIVMLQINPDS 446
 DB 415 PKDLKRLMVFRRGEELDYGVAREMLYLLGCEMNPYGYLFQYSTDNIVMLQINPDS 474
 QY 447 INPDLSYFHFVGRIMGLAVFHHYINGFTVPYKQLGKPIQLSDLESVDELKSLV 506
 DB 475 VNPEHLSYFHFVGRIMGLAVFHHYINGFTVPYKQLGKPIQLSDLESVDELKSLV 534
 QY 507 WLENDITGVLDHFTVCENHAYGIIQHELEKPKGSI PVNEENKSEYVRLYNNRFLRGI 566
 DB 535 WLENDITGVLDHFTVCENHAYGIIQHELEKPKGSI PVNEENKSEYVRLYNNRFLRGI 594
 QY 567 BAQFLALQKGFNEILPOHLKPPDQKELIELIGLDKIDLDWKSNTRLKCVASNIYR 626
 DB 595 BAQFLALQKGFNEILPOHLKPPDQKELIELIGLDKIDLDWKSNTRLKCVASNIYR 654
 QY 627 WFWQAVRFPDEERARLLOPTGSTRVPLQGFALQGSTAAGPRPLTHLIDANTDNL 686
 DB 655 WFWQAVRFPDEERARLLOPTGSTRVPLQGFALQGSTAAGPRPLTHLIDANTDNL 711
 QY 687 KAHTCFNRIDIPYSEYEKYEKLLTAVETCGPAVE 723
 DB 712 KAHTCFNRIDIPYSEYEKYEKLLTAVETCGPAVE 748
 XX
 XX RESULT 6
 XX AAU19610
 XX ID AAU19610 standard; Protein; 804 AA.
 XX
 XX AAU19610;
 XX
 XX 04-DEC-2001 (first entry)
 XX
 XX Human diagnostic and therapeutic polypeptide (DITHP) #196.
 XX
 XX Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
 XX cell proliferative disorder; Crohn's disease; lymphoma; leukemia;
 XX acquired immune deficiency syndrome; AIDS; autoimmune disorder;
 XX respiratory disorder.
 XX
 XX Homo sapiens.
 XX
 XX MO200162927-A2.
 XX
 XX 30-AUG-2001.
 XX
 XX 21-FEB-2001; 2001WO-US06059.
 XX
 XX 24-FEB-2000; 2000US-0184693.
 XX 24-FEB-2000; 2000US-0184697.
 XX 24-FEB-2000; 2000US-0184698.
 XX 24-FEB-2000; 2000US-0184768.
 XX 24-FEB-2000; 2000US-0184769.
 XX 24-FEB-2000; 2000US-0184770.
 XX 24-FEB-2000; 2000US-0184771.
 XX 24-FEB-2000; 2000US-0184772.
 XX 24-FEB-2000; 2000US-0184773.
 XX 24-FEB-2000; 2000US-0184774.
 XX 24-FEB-2000; 2000US-0184776.
 XX 24-FEB-2000; 2000US-0184777.
 XX 24-FEB-2000; 2000US-0184779.
 XX 24-FEB-2000; 2000US-0184813.
 XX 24-FEB-2000; 2000US-0184817.
 XX 24-FEB-2000; 2000US-0184841.
 XX 24-FEB-2000; 2000US-0185213.
 XX 24-FEB-2000; 2000US-0185216.
 XX 12-MAY-2000; 2000US-0203785.
 XX 15-MAY-2000; 2000US-0204226.
 XX 16-MAY-2000; 2000US-0204525.
 XX 16-MAY-2000; 2000US-0204821.
 XX 16-MAY-2000; 2000US-0204908.

PR 16-MAY-2000; 2000US-0205232.
 PR 17-MAY-2000; 2000US-0204815.
 PR 17-MAY-2000; 2000US-0204863.
 PR 17-MAY-2000; 2000US-0205221.
 PR 17-MAY-2000; 2000US-0205285.
 PR 17-MAY-2000; 2000US-0205286.
 PR 17-MAY-2000; 2000US-0205287.
 PR 17-MAY-2000; 2000US-0205323.
 PR 17-MAY-2000; 2000US-0205324.
 PA (INCYTE GENOMICS INC.)
 XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
 PI Chen A, D'Sa SA, Amstey S, Dahl CR, Dam TC, Daniels SE,
 PI Dufour GR, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL,
 PI Liu TF, Roseberry AM, Russo BH, Russo FD, Stockreher TK, Daffo A;
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W,
 PI Cohen HU, Hodgson DM, Lincoln SE, Jackson S;
 XX WPI; 2001-502867/55.
 DR N-PSDB; AAS31181.
 XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics
 PT -
 PS Claim 27; Page 512-514; 522pp; English.
 XX The invention relates to polynucleotides (I) encoding diagnostic and
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,
 CC and proteins involved in growth and development and receptors. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate DITHP expression. For example, (I) and
 CC (II) may be used to treat disorders associated with decreased polypeptide
 CC expression by rectifying mutations or deletions in a patient's genome,
 CC that affect the activity of the DITHPs, by expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II)
 CC may be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of
 CC antibodies against DITHPs and in assays to identify modulators of DITHP
 CC expression and activity. The anti-DITHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DITHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbant
 CC assay (ELISA)). AAU19415-AAU19625 represent human diagnostic and
 CC therapeutic (DITHP) polypeptides of the invention.
 XX
 XX Sequence 804 AA;
 SQ
 Query Match 77.1%; Score 2996; DB 22; Length 804;
 Best Local Similarity 74.5%; Pred. No. 2.3e-285;
 Matches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;
 QY 5 IKIRLTVCANLKKKQDFRLPDPFAKIYVDSGQSGSTVTNKKTDPPKNNHGYDLYVKK 64
 DB 68 VKHLTLTVCAKLVKKQDFRLPDPFAKIYVDSGQSGSTVTNKKTDPPKNNHGYDLYVKK 127
 QY 65 TDSITISVWNHKKHKHKKAGAGFLCCVRLSNALSRLKDTGYQRLDLCKLNPSTDAVRGQ 124
 DB 128 SDSVTISVWNHKKHKHKKAGAGFLCCVRLSNALSRLKDTGYQRLDLCKLNPSTDAVRGQ 187
 QY 125 IVVSLQTRDRIGTGGSVYDRCGLLEN-----GTV----- 155
 DB 188 IVVSLQSRDRIGTGGSVYDRCGLLEN-----GTV----- 247
 QY 156 --EDSGPGRPLSCFMEBPAPYTDSTGAAGGNGCRFVESPSQDRLQAOQLRNPVGRSL 213

DB 248 ASYSSPGRPLSCFMEBPAPYTDSTGAAGGNGCRFVESPSQDRLQAOQLRNPVGRSL 298
 QY 214 QTPQNRPHGQSPPELPEGEBOBTVOGVYFLATOTGVSFWHDPRIIPRINSVNCDELGP 273
 DB 299 ---SKTHLHTPDDLPEGEBOBTVOGVYFLATOTGVSFWHDPRIIPRINSVNCDELGP 354
 QY 274 LPPGWEKSTVSGRIYFVDHNNRTQFTDPR---LHHNNHQQQLKEPQPLPPEEGS 329
 DB 355 LPPGWEKSTVSGRIYFVDHNNRTQFTDPR---LHHNNHQQQLKEPQPLPPEEGS 410
 QY 330 L---EDDELPACQERDLYVQKLYRHLSTQQPQAGCRTEVERRETFEESYQIMKR 386
 DB 411 LCPDDTECLTVPRYKRDLYVQKLYRHLSTQQPQAGCRTEVERRETFEESYQIMKR 470
 QY 387 PKDLKRLMVKRGEGBLDYGVAREWLYLCHENLNYPGLFOYSTDNIMYLQINPDS 446
 DB 471 PKDLKRLMVKRGEGBLDYGVAREWLYLCHENLNYPGLFOYSTDNIMYLQINPDS 530
 QY 447 INPDHSTYHFGRIYGLAVFHHGTHNGFTVPPKOLLGPIQLSDLESVPDELHSTV 506
 DB 531 VNEHLSYHFGRIYGLAVFHHGTHNGFTVPPKOLLGPIQLSDLESVPDELHSTV 590
 QY 507 WLENDITFVLDHTFCVEHNAFGRIYGLAVFHHGTHNGFTVPPKOLLGPIQLSDLESVP 566
 DB 591 WLENDITFVLDHTFCVEHNAFGRIYGLAVFHHGTHNGFTVPPKOLLGPIQLSDLESVP 650
 QY 567 EAQFLALQKGFNEVYHQLLKPDDKELELLIGLKDLDMDKSNFTLKVQVADSNIVR 626
 DB 651 EAQFLALQKGFNEVYHQLLKPDDKELELLIGLKDLDMDKSNFTLKVQVADSNIVR 710
 QY 627 WFOAVETDEERRARLLQFVTSRVPVLOGFKALQSGTGAAPRLFTIHLIDANTDNL 686
 DB 711 WFOAVETDEERRARLLQFVTSRVPVLOGFKALQSGTGAAPRLFTIHLIDANTDNL 767
 QY 687 KATGCRNRIDIPYSEYKLYKLLTAVETGPAVE 723
 DB 768 KATGCRNRIDIPYSEYKLYKLLTAVETGPAVE 804
 RESULT 7
 ID AAU87301 standard; Protein; 514 AA.
 XX AAU87301;
 AC
 AC
 DT 05-JUN-2002 (first entry)
 XX
 DE Novel central nervous system protein #211.
 XX
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischemia; angioneurosis;
 KW nervous system disorder; Alzheimer's disease; gastrointestinal disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminization;
 KW endocrine disorder; diabetes; cancer; leukemia; neovascularization;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food preservative; gene therapy.
 OS Homo sapiens.
 XX
 XX W0200155318-A2.
 PD 02-AUG-2001.
 XX
 PD 17-JAN-2001; 2001MO-US01332.
 PF
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.

CC The invention describes an isolated nucleic acid molecule (I) encoding a
 CC novel central nervous system protein. (II) and polypeptides (III) encoded
 CC by (I), are used to treat a medical conditions and in diagnosis of a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. Rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 CC adenocarcinomas and irritable bowel syndrome, reproductive system
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 CC leukemia, disorders involving neovascularisation e.g. malignancies,
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 57.6%; Score 2236; DB 22; Length 514;
 Best Local Similarity 80.7%; Pred. No. 8.8e-211;
 Matches 414; Conservative 43; Mismatches 42; Indels 14; Gaps 4;

QY 218 NRPHGOSPELPGYGEORTVQGVYELHTQGVSTWHPERIPROLNSVNCDEGLPRLG 277
 DB 9 SRTHLTPPLPBGYGEORTVQGVYELHTQGVSTWHPERIPROLNSVNCDEGLPRLG 68
 QY 278 WEVASTVSGRIYVDNNRTTFTDPR---LHIMNHOCQKESQPLPSSGSL--- 330
 DB 69 WEIRMTATGRVYFVDHNNRTTFTDPRLSANHLVLRNQKLDQCCQGV---SLCPD 124
 QY 331 EDELPAGQRYERPLVOKLKYRLHSLQOPQAGHCIEVSREIFEESSROIMKWRPDL 390
 DB 125 DTECLTVPRYKRDVQKTLRQELSQOPQAGHCIEVSREIFEESSROIMKWRPDL 184
 QY 391 KKRIMVFRGEEGLDYGVARWLYLICHEMLNPYYGLFQYSTDNIMQLNPSSIND 450
 DB 185 WKRLMKFRGEEGLDYGVARWLYLICHEMLNPYYGLFQYSTDNIMQLNPSSIND 244
 QY 451 HLSYFHFVGRIMGAVHGHYINGFTVPRYKOLGKPIQLSDLSVPELHKSLWTLE 510
 DB 245 HLSYFHFVGRIMGAVHGHYIDGFTLPPYKOLGKSLTLDMELVDDLEHNSLWTLLE 304
 QY 511 NDITPVLDRHFCVEHNAFGRILCHEKENGKRVNVTENKESYRLVYVWRFRGIEAQF 570
 DB 305 NDITGVLDHFCVEHNAFGEIICHEKPNKGSIPVNEENKEVYRLVYVWRFRGIEAQF 364
 QY 571 LALQGFNELIPQHLKRPDQKLELITGLDKIDLNKSNTRLKCYADSNITVYRWQ 630
 DB 365 LALQGFNEVIPQHLKRPDQKLELITGLDKIDLNKSNTRLKCYADSNITVYRWQ 424
 QY 631 AVETFEDEERARLLQFTVGSSTRVPLQGFALQSGSTAGAPRLFTIHLIDANTDLPKAT 690
 DB 425 AVETFEDEERARLLQFTVGSSTRVPLQGFALQ---GAAPRLFTIHLIDANTDLPKAT 481
 QY 691 CFNRIDIPYESTSEKLYEKLTAVEETCGFAVE 723
 DB 482 CFNRIDIPYESTSEKLYEKLTAVEETCGFAVE 514

RESULT 8
 ABB61120
 ID ABB61120 standard; Protein; 1035 AA.
 AC ABB61120;
 XX
 XX
 DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster polypeptide SEQ ID NO 10152.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN MO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL05223.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 10152; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB120511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB85737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 1035 AA;

Query Match 53.3%; Score 2069; DB 22; Length 1035;

Best Local Similarity 41.2%; Pred. No. 8.6e-194;
 Matches 435; Conservative 104; Mismatches 145; Indels 372; Gaps 15;

QY 6 KIRLTVLCAKNIAKDFRLPDPFAKIVVDGSGCHSTDTVYKNTLDRKMQHDLVYGT 65
 DB 14 KVRITTLCAKNIAKDFRLPDPFAKIVVDGSGCHSTDTVYKNTLDRKMQHDLVYGT 73
 QY 66 DSITISVNHKKIKHKKQAGFLGCVRLSNASIRLKTGYORLDCIKLNPSTDAVVGQI 125
 DB 74 DATITVNMQRKIKH--GSGFLGCVRIAFNFIQSLKAGFRRLDGLTSPDDDLVAVGQI 131
 QY 126 VVSLQTRDRIGT-----GSVYDCRGLIENE-----GTVY----- 155
 DB 132 ITSLSKDPSGPNPLATVPGSDVGRGSEDDSSBDSLPEGMEERRINDNGRVYVNHATK 191
 QY 156 ----- 155
 DB 192 STQWDRPRPGVVGSSHATSPQQRNTNNGSGDRQAPAGTRSTTCTNLMNNHRSRL 251
 QY 156 -----EDSGGRPLSCFMEBPADYDSTAAAGGNCRFV----- 191
 DB 252 SVTASDERHSTELISVYKENTSTPTVSA--TTTPCKKTSNSSSSSGG--RTLEGRPT 308
 QY 192 ----SPSQQRLOAQRLNPNVYRGLTPORRPFQHSQE----- 227
 DB 309 NEPATPTSTTSASVRLHNSD--NHVKIPKQCTGNHAPESTPTSPGCGQNVYVNGANG 366
 QY 228 ----- 227

```

Db 367 STSGNGSGQAQPOSASNGWTQEDAAITTSPTTSSPPRHSGSPPTINISPAATPSAN 426
Qy 228 ----- 227
Db 427 GNVSFMANSTPAGSGGGSSYTAATPGQRSGRSSRQGEESSTRRSRGTNGCTSG 486
Qy 228 -----LPEGEQRTTVQGVYFLHQTGVTWHDRI 259
Db 487 GGGGGGGGQRYASAAIAAQAAPFLIDLPFGEMKTTQGGGVYFHTIGVSWHDPRI 546
Qy 260 PRDINS--VNCDELGPUPGMEVASTVSGRIYFVDHNNRTTQFTDPRLLHIM----- 309
Db 547 PRDFTQHLTLDAGPLPSGMEGRKTAAGKIVYVDHNNRTTQFTDPRLLSGSLIQIRRG 606
Qy 310 ---NHQCLKPSQPLPLPS-----EG 328
Db 607 VPTSAANAAGTPAPPSATPATPSAAAAPPOATPASNAAPPTLTITNDPARIIVDLPOG 666
Qy 329 SLEDELPAQRRELDVQKLKVLNHELSLQPOAGHCRLEVSREEIFEESYRQIMKMEPK 388
Db 667 LLEGADL-LPKYRDLVQKLRALRTELQTMQPSGHCRLVSSNEIFEESYRLIMKMAK 725
Qy 389 DLKRLMYKFRGEGGLDYGVAREMVLTLCHENLNPYGLFOYSTDNITMLQINDSSIN 448
Db 726 DMKRLMYKFRGEGGLDYGVAREMVLTLCHENLNPYGLFOYSTDNITMLQINDSSIN 785
Qy 449 PDHLSYFHFVGRINGLAVFHGHYINGGFTVPFYKQLGKQLSDLESVDPELHKSILWI 508
Db 786 PDHLSYFHF-----LINRPTIGLDEGVDPDLHSLTWM 819
Qy 509 LENDITPVLDHTFCVENNAFGRLLQHELKNGSNVPTEENKKEYRVLVYNNRPMGIEA 568
Db 820 LESNISGIIESTFSEVNNRFGALVHLELKDGASIPTEENKKEYRVLVYNNRPMGIEA 879
Qy 569 QELAQGFENLLOHLLKPKPDQKELIIGGDKLIDNWKSNTEKCVADSNVYRF 628
Db 880 QFLAQGFCELIPLSHLRPFDERELELVIGLSSIDVNDNRNNTLKHCTVETQVLMF 939
Qy 629 WQAVETPDEERRARLLQFTVGTSTVPLQGFALQSGTGAAPRLFTIH-IDAINDLPLK 687
Db 940 WQAVESYSEMRARLLQFTVGTSSAVPLQGFALQSGTGAAPRLFTIHLDVPTGNLXK 999
Qy 688 AHTCFNRIDIPRESYEKLYEKILTAVEETCGRAVE 723
Db 1000 AHTCFNRIDIPRETYQLCDKLQWAVEETCGRAVE 1035

RESULT 9
AAM78877
ID AAM78877 standard; Protein, 335 AA.
AC AAM78877;
XX 06-NOV-2001 (first entry)
XX 06-NOV-2001 (first entry)
XX Human protein SEQ ID NO 1539.
DE Human protein SEQ ID NO 1539.
XX Human, cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.
XX Homo sapiens.
XX MO200157130-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001MO-US04098.
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.

```

```

PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang Y, Ren F, Chen R, Wang ZW;
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
XX DR N-PSDB; AAK52010.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX PT useful in diagnosis and gene therapy -
XX Claim 20; Page 3847; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activity/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52591), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX Sequence 335 AA:
SQ
Qy 168 MEEPAVPTDSTGAAGGNCRFYESSPODRLQAOQLNRPDVGSLQTPORPHGOSPE 227
Db 1 MEEPAVPTDSTGAAGGNCRFYESSPODRLQAOQLNRPDVGSLQTPORPHGOSPE 60
Qy 228 LPGEYEQRTTVQGVYFLHTQGVSTWHDRIIPRLINSVNCDELGLPFGMEVASTVSGR 287
Db 61 LPGEYEQRTTVQGVYFLHTQGVSTWHDRIIPRLINSVNCDELGLPFGMEVASTVSGR 120
Qy 288 IYVDHNNRTTQFTDPRLLHINNHOCLKEPSQPLPLBSGSLDEELPAQRYERDLYQK 347
Db 121 IYVDHNNRTTQFTDPRLLHINNHOCLKEPSQPLPLBSGSLDEELPAQRYERDLYQK 180
Qy 348 LKVLRLHLSLQCPQAGHCRLEVSREEIFEESYRQIMKRPDLKRLMYKFRGEGGLDYG 407
Db 181 LKVLRLHLSLQCPQAGHCRLEVSREEIFEESYRQIMKRPDLKRLMYKFRGEGGLDYG 240
Qy 408 GVAREMVLTLCHENLNPYGLFOYSTDNITMYLQINDSSINDHLSYHFGRINGLAVF 467
Db 241 GVAREMVLTLCHENLNPYGLFOYSTDNITMYLQINDSSINDHLSYHFGRINGLAVF 298
Qy 468 HGHYINGFTVPFYKQLGKQLSDLESVDPELHKSILWIENDITVLDHTFC 522
Db 299 -----CVFWTLHQRG-----LHSLALLQAAAGAHPLRSGIC 330

RESULT 10
AAM13385
ID AAM13385 standard; Protein, 766 AA.
AC AAM13385;
XX 10-JUL-1997 (first entry)

```

XX DE Human protein ubiquitin ligase pub2.
 XX KM Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase;
 XX KM p53; cell cycle; transgenic animal.
 XX OS Homo sapiens.
 XX PN MO9712962-AL.
 XX PD 10-APR-1997.
 XX PF 04-OCT-1996; 96MO-US15930.
 XX PR 04-OCT-1995; 95US-0539205.
 XX PA (COLD-) COLD SPRING HARBOR LAB.
 XX PI Beach D, Caligiuri M, Nefsky B;
 XX DR WPI; 1997-226206/20.
 XX DR N-PSDB; AAT47041.
 XX PT Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved
 PT in cdc25 phosphatase and p53 ubiquitination, and regulate cell
 PT growth and proliferation
 XX PS Claim 1; Page 80-84; 108pp; English.
 XX CC Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and
 CC pub3 (AAW13386) are homologues of fission yeast pub1 (AAW13387) and
 CC were identified from cDNA clones (AAT47040-42) obt'd. e.g. from a
 CC keratinocyte cDNA library. Pub polypeptides can be produced in
 CC transfected host cells. They can control the steady state level of
 CC cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2)
 CC dephosphorylation and the steady state level of p53 (controlling
 CC the degree of cell cycle regulation of p53). They can regulate
 CC cell or tissue differentiation, or cell growth or proliferation by
 CC affecting other proteins, can be a specific (ant)agonist of wild-
 CC type protein function and may be used as immunogens to elicit a
 CC specific immune residue.
 XX SQ Sequence 766 AA;
 Query Match 37.6%; Score 1461; DB 18; Length 766;
 Best Local Similarity 40.8%; Pred. No. 4.5e-134;
 Matches 326; Conservative 118; Mismatches 236; Indels 120; Gaps 20;
 QY 4 SIKRLVLCANLAKDPFRLPDPFAKIVVDSGQCHSTIVKNTLDPKMGHYDLYVG 63
 DB 7 SRIKIVIVAAADGKYKRDVFRFPDPFAVLTVDGE-QTHTTAIKKTILNPFYMNTEFEVNT 65
 QY 64 KTDSTISVMNHKKIKHKGAGFLGCVRLSNALSKDGYQL--DLCKNPSPTDLY 121
 DB 66 DNSTAIQVPOCKE-KKGGGFLGVINLVGDVLDLIGDEMLICDLKLN--ENTTV 122
 QY 122 RQGIIVSLQ-----TRDRIGTGSVVDORGLLENEGTYED 157
 DB 123 HGKIIINLSTAGQLTVPSAASGARTORTSINDPOSSKSSSVSNPASAGPTBD 182
 QY 158 SGPG-----RPLSCMEB---PAY---TSTG-----AAA 182
 DB 183 NAPAASPASSEPRFTSSFDQGRLPGEWRCITNLRITVVDHNTKSTWIRPNLSVA 242
 QY 183 GGGNCRFVESP-----ODRLOAQRLENPVRSGLQTPQNRPHGOSPELPFGVYQ 234
 DB 243 GAAAEHLSSASAVTEGVQPSNNARTEAVALSNMTAG-----SGELLGWGQ 296
 QY 235 RTTVGGVYFLHTQGVSTWMDPRIPRDLNSV-----CDELGLPFGMEVRS 283
 DB 297 RYTPGGRPYFVDHNRRTTWDPRRQGYIRSYGGPNNAITQQQFVSGLPFGWEMRLT 356
 QY 284 VSGRIYFVDHNRRTQFTDPRILHIMNQCLKEPSQPLPSESGLEDELPQGYERD 343

DB 357 NTARVYFVDHNRKTTTMDPR-----LPS--SL-DQVVP--QYKRD 392
 QY 344 LVOKKLVLEHLSLOQPOAGCRIVSREIFPESYQIMMRKDLKRLMVFRRBEG 403
 DB 393 FRKKLYFLSQPAL-HPLPGQCHIVRNNHIFEDSYAIMEQATDLKRLMTFDSBG 451
 QY 404 LDYGVAREMLYLICHEMLNPFYGLFOYSTNIVMLQINPDSINPHLSYFHFVGRIMG 463
 DB 452 LDYGLSRIFYFLSHENMNPYCLFEVSSVDNTLQINPDSINPHELVNFKIGVIG 511
 QY 464 LAVFHGYINGFTVPFYKQLGKPIQLSDLESVDPELHKLWILENDITPVLDHPCV 523
 DB 512 LAIFHRFVDAFVFFVYEMILQCKYTLQDMESMDAEYRSLWILNDIGVLDLFSV 571
 QY 524 ENHAFGRILQHEIKPGRVVPVTEENKKEVYRLVYVFMFGIAQFATLOKGFNEIPQ 583
 DB 572 EDNCFSEVTTIDKRGRIEVTENKEVYDLYVW-IQRIEGRFAHFGSSELLPQ 630
 QY 584 HLKPPDQKELELIIGLDKIDLNDWKSNTRLKRCVADSNIVRMFQAVETFEERRARL 643
 DB 631 ELINVDERELELIGLSEIMEDMKCKDYRSYSENDQIKFMELMDWSENKXSRL 690
 QY 644 LQFTGSTVPLQGFYALQSGTAGAPRLFTHTLIDANTDLPRXHTCFNRIDIPPEY 703
 DB 691 LQFTTGISRLPVNGFKDLOGSD--GPKFTIERK-GEPNKLPRAHTCFNRIDIPYTSK 746
 QY 704 EKLVEKLTAVEETCGFAVE 723
 DB 747 KDLDHKLSTAVEETIGFGE 766
 RESULT 11
 ID ABP73459 standard; Protein; 832 AA.
 XX ABP73459;
 AC ABP73459;
 XX 30-JAN-2003 (first entry)
 XX
 DE Candida albicans essential protein SEQ ID NO 7296.
 XX
 XX Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 KW signal transduction; DNA replication; cell division; growth;
 KW proliferation; Candida albicans; fungicide; antifungal.
 OS Candida albicans.
 PN WO200253728-A2.
 PD 11-JUL-2002.
 PF 26-DEC-2001; 2001MO-US49486.
 PR 29-DEC-2000; 2000US-259128P.
 PR 20-FEB-2001; 2001US-0792024.
 PR 22-AUG-2001; 2001US-314050P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KU;
 XX WPI; 2002-566694/60.
 DR N-PSDB; ABZ3209.
 XX Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele
 PT of a gene and placing other allele of the gene under conditional
 PT expression -
 PS Claim 44; SEQ ID NO 7296; 167bp + Sequence Listing; English.
 XX The invention relates to constructing (M1) a strain of diploid fungal

cells in which both alleles of a gene are modified, comprising modifying one allele by insertion or replacement by a cassette having an expressible selectable marker and modifying other allele by recombination, of a promoter replacement fragment with a heterologous promoter, so that expression of the second allele is regulated by the promoter. (M1) is useful for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The diploid fungal cells having both alleles modified are useful for identifying a gene that is essential to the survival or growth of a fungus, a gene that contributes to the virulence and/or pathogenicity of a fungus, a gene that contributes to the resistance of a diploid fungus to an antifungal agent, an antifungal agent that inhibits the growth of a diploid fungus and for identifying a therapeutic agent for treatment of a mammalian disease. (M1) is useful for identifying a compound which modulates the activity of a gene product, preferably enzymatic activity, carbon compound catabolism, biosynthetic, transporter, transcriptional, translational, signal transduction, DNA replication and cell division activity. The method is useful for identifying a compound having the ability to inhibit growth or proliferation of *C. albicans* cells and for treating infection by *C. albicans*. The present sequence is that of an essential *Candida albicans* protein used in the method of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office.

Sequence 832 AA;

Query Match 37.0%; Score 1436.5; DB 23; Length 832;

Best Local Similarity 36.4%; Pred. No. 1.3e-11;

Matches 326; Conservative 124; Mismatches 233; Indels 167; Gaps 21;

7 RLTVLCAKNIKKDFFRLPDPFAKIVDVGSGCHSTDTYKNTLDPRKNGHYDLYGKTD 66
 17 INVKVALESYKRDVFRQDPDFAVLTVDGS-QTKTITAKKTINPYMNETFN-FOAKED 74
 67 SI-TISVNNHKKIKHKGAGFLGCVRL-LSVAISRLMDTGVRLL-DLCLNDSDDAR 122
 75 SILVIVQDFDKKF-KKQDGFVGVINRIGVIDLINSSEETITRLKSN--ENLAAS 131
 123 GOIVVSLQ-----TRD-----RIGTGSVVDCKGLLE 149
 132 GKIIVISHNNSNGGVTATTTGTGASSNNIATITSGVNNLRISATTTANSTQAAS 191
 150 NEGTV-YEDSGPCR--PLSCMEEPAYTSTGAAGGCKRVESPSDDQLQQRRL 205
 192 SPATVANGSGPTSLPPIGQGHPEPTATPGAGAAAGASROYSPEDQYGLPQWER 251
 206 NPDVAG-----SLOTPQNRPHCHQSP----- 226
 252 RTDNFGRITYVDHNRITTTQGRPALHQSERTERGQROSEFBARRQKRTLDGEGSV 311
 227 -----ELPEGYEQRTVQGV 242
 312 LPTGSGNSITSGNTVNASGANTPVNPAAYMMAAGATTSGLDELPSGEGORFTTEGR 371
 243 YELHQTGVSTWHDRIPIRDLNSVN-----CDELGPLPQMEYRSTVSGIYVCH 293
 372 YVDHNRITTTWVDRRQGYRTFPGPNTTIOQPVSGISGIPSPSEMERLNTARVIVDH 421
 294 NNRTQFTDPRLHIMNHQCOLKEPSQPLPSPSEGLEDEELPAQRERDLVQKLKYLRL 353
 432 NTKTTWDDPR-----LPS--SL-DQVNP--QYKRPFRFVYIFRS 467
 354 ELSLQPGQGHCRIVSREEIFEESYQIMKRPYDKLKKRLMVFREBEGLDGVGVAREW 413
 468 QPAL-RILPGQCHIKVRDHIFEDYQIMKOTPEBCLKRLMIFDGBEGLDGVGVAREW 526
 414 LYLCHENLNPYGLFOYSTDNIVMLQINPDSINPDLSYFHFVGRIMGLAVHGYIN 473
 527 PFLSHDMENPFYCLFEVSSHNDNYTLQINPDSINPEHLNFKIGRVGLGVHRRFLD 586
 474 GGFVDFYKQLGKPIQLSDLESVDPELAKSLVWLTENDIPVLDHFFVEHNAFGRILO 533

DB 587 AFVFGALYKMLHKVYVLDMEGVDAEFRLSKMLINDITGILDITSABEESGZIVE 646
 QY 534 HEKENGNAVPTTEENKKEYVRLVYWRPMEGIEQFLALQGFNELLIPOLHLPDQKE 593
 DB 647 VDLKPGGRDIEVEENKHEVELITEMRISKVEEQFKAFIFGFNELLIPOLVYVDFRE 706
 QY 594 LELIGGLDIDLDIMKSNTRLGHCVADSNIVRMPQAVETFEERRRARILOFTVSTRV 653
 DB 707 LELIGGLAIDEDMKKHDYRGRVQENQVIGVWKICINENDSOKARLQFTTGTSRI 766
 QY 654 PLOGFKALQSGTGAAGPRFTIHLIDANTDNLPKATCFNRIDIPPYSEYKLYKLLTA 713
 DB 767 PVNGFKDLQSSD---GPRFTIEKA-GEANQPKSHTCFNRVLDLPYDYSIKQKTLTA 822
 QY 714 VERTGPAVE 723
 DB 823 VERTVGGQE 832

RESULT 12

ABJ26104 standard; Protein; 869 AA.

ABJ26104;

16-APR-2003 (first entry)

Aspergillus fumigatus essential gene protein #762.

Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection; cancer; contamination; biofilm; antibody; immune response.

Aspergillus fumigatus.

MO200286090-A2.

31-OCT-2002.

23-APR-2002; 2002WC-US13142.

23-APR-2001; 2001US-285697P.

27-APR-2001; 2001US-287066P.

05-JUN-2001; 2001US-295890P.

09-JUL-2001; 2001US-303899P.

31-AUG-2001; 2001US-316362P.

(ELIT-) ELITRA PHARM INC.

Jiang B, Tiehoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,

WPI, 2003-093124/08.

New purified or isolated nucleic acids of essential genes of *Aspergillus fumigatus*, useful for treating or preventing infections by *A. fumigatus*, or for treating a non-infectious disease in a subject e.g. cancer -

Disclosure: Page -: 175pp, English.

The invention relates to novel purified or isolated nucleic acids of essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as *A. fumigatus*, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by *A. fumigatus*, or to prevent or inhibit formation on a surface of a biofilm comprising *A. fumigatus*. The polynucleotides are useful for expressing recombinant protein for characterization, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of *A. fumigatus* to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and

CC invention.
 XX Sequence 854 AA;
 SQ Query Match 33.9%; Score 1318; DB 20; Length 854;
 Best Local Similarity 35.4%; Pred. No. 6.8e-120;
 Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

QY 6 KIRLVLCAR-NLAKDFFLLPDPFAKIYVDGCGCHSDYTKNTLDPMMCHNLYLXG 64
 DB 9 QLOTVISATLKKKNKXWFG-PSPYVEVVD--GQSKTEKCNNTNSPKWKQPLTVITP 65
 QY 65 TDSITISVNNKKIKKKQAGFLGCVRL-----LSNAISRLKDTGYCR----- 107
 DB 66 TSKLCFRVWGHQTL--KSDVLLGTAGLDIYETLKSNNKLEEVVMTQLVADKPEFTM 122
 QY 108 --LDLCKLPSDDIVANGQIVS-----LQTD--RIGTGS--VDCR 145
 DB 123 GDLSTVC---LDGLQVAAEVVINGETSCESTTQNDGCRITDDTRVSTNGSEDEVAS 178
 QY 146 G-----LLENQGVYEDSGGRPLSCFMEBPAYT-----DSTG 179
 DB 179 GENKRGANNSPSLNSG--FKESRPRP-----SRPPPTRRPASVNGSPSTSDSDG 231
 QY 180 AA-----AGGNCRFVESPSQ-----DQRLQAR 203
 DB 232 SSTSLPPTNTVNTSTEGATGGLIPLTISGSGSPRLNTVSCAPLPPGWEQKRV--- 287
 QY 204 LRNPVAGSL-----QTPONRPHGHOSPELPEGYBORTVQGOVYFLHTQGVSTW 255
 DB 288 ---DQGRVYVVDHVEKRTTWDRP-----PLPGEKRVDMKRIYVVDHFTRTTWQ 338
 QY 256 DPRIP-----BDL-----NSVNCDELGLPPGMEVR 281
 DB 339 RPTLESVRNVEQWOLQSQAMQCFNQRFYIGNODLPATSONKEFDLGLPPGMEKR 398
 QY 282 STVGRIYFDVHNHRTQPTDPRLHNMHCCKLSPQPLP----- 324
 DB 399 TDSNGRVYFVNNHRTIQWEDPR-----SQQLNE--KPLPGEMERTVDCIYFVDH 450
 QY 325 -----PSEGLEDEELPAQYERDLVQKLVR-----HELSLQCPAGHCRIEVSRE 372
 DB 451 NRRATYIDPRTGKALDNGQIAYVDFKAKYQYFRFMOQLAMPQ---HIKITVTRK 506
 QY 373 EIFESYRQIMKMKPKLKKRLMKFEGEGLDYGVABEMVLLCHEMLNYYGLFOFS 432
 DB 507 TLFEDSFQOIMSFSPQDLRRRLMWIFPGEGLDYGVAARWFFLLSHEVLNPMYCLFEYA 566
 QY 433 TDNIYMLQINPDSSINDHLSEFHFVGRIMGIAVFHGYINGFTVPFYKQDLGKPIQS 492
 DB 567 GKDNVYCIQINPASTINDHLKTPFRIGRTAMALFPGKFLIDGFSLPFYKRLINKPVGLK 626
 QY 493 DLESDVPELHKSLLVILNDITPV-LDHTCVBNAFGRLLQHELKPNNGNVPTEENK 551
 DB 627 DLESIDPEFYNLSLWVENNIEECGEMYSVXKEILGKSHDLKPNQNLVTEENK 686
 QY 552 EYVRLVYVNRFMRGIEQFLALQKGFNELPOHLKPFQDKEHLITIGLDKIDNDMS 611
 DB 687 EYIMVAMERLSRVEGQTAFFEGFRIILPQOYLQFDKKELEVLICMGEILDNDMR 746
 QY 612 NTRLKGVADSNIVRFWQAVETFDERRARLLQFVTGSTRVLPQGFKALQSGTGAAGR 671
 DB 747 HAIRHAYRTSKQIMFWQFKEIDNEKEMRLIQFVGTGCELRVGGFADIMSGN--GPO 803
 QY 672 LFTHLIDANTDNLPKAHTCFNRIDIPYASYELVEXKLITAVEFCGFAVE 723
 DB 804 KFCIEKV-GKENMLPRSHTCNRLDLPYASYELKELILFAIEFGGQ 854

AC AAY0948;
 XX 21-OCT-1999 (first entry)
 DE Human E3 ubiquitin protein ligase protein.
 XX E3 ubiquitin protein ligase; h-E3 UPL; antiinflammatory; antidiabetic;
 KW immunosuppressive; neuroprotective; cyostatic; antiarthritic; cardiant;
 KW immunomodulator; antiviral; treatment; screening; gene therapy; cancer;
 KW inflammation; autoimmune disease; neurological disease; apoptosis;
 KW endothelial cell; proliferation; differentiation; angiogenesis; cachexia;
 KW peripheral vascular disease; hematopoietic disorder; arthritis; leukemia;
 KW pulmonary disorder; diabetes; viral infection; human.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 275..306
 FT /label= WWI
 FT /note= "WW protein interaction domain I"
 FT Domain 307..340
 FT /label= WWII
 FT /note= "WW protein interaction domain II"
 FT Domain 386..420
 FT /label= WWIII
 FT /note= "WW protein interaction domain III"
 FT Domain 427..460
 FT /label= WWIV
 FT /note= "WW protein interaction domain IV"
 FN NO9940201-AI.
 XX 12-AUG-1999.
 PD 02-FEB-1999; 99WO-GB00353.
 PF
 XX 30-APR-1998; 98US-0070060.
 PR 05-FEB-1998; 98US-0073839.
 PA (ZENB) ZENBICA LTD.
 XX Ghildyal N, Husted CM;
 PI WPI: 1999-508506/42.
 DR N-PSDB; AAY09235, AAY09236.
 DR
 XX New human proteolytic accessory enzyme and its modulators useful
 PT for treating disease conditions like inflammation or autoimmune
 PT diseases
 PS Claim 1; Fig 3; 95pp; English.
 XX
 CC This invention describes a novel human polynucleotide (I) which encodes a
 CC E3 ubiquitin protein ligase, h-E3 UPL (II). The products of the invention
 CC have antiinflammatory, immunosuppressive, neuroprotective, cyostatic,
 CC antiarthritic, immunomodulator, antidiabetic, antiviral and cardiant
 CC activity. The products of the invention can also be used for treating
 CC patients with disorders mediated by the biological and/or pharmacological
 CC activity of h-E3 UPL. The nucleic acid sequences encoding E3 UPL are used
 CC in expression systems as assay for agonists and antagonists for the E3
 CC UPL protein. The E3 UPL protein is used in screening assays to identify
 CC blockers and antagonists. They are also used in gene therapy. Specific
 CC modulation of biological and/or pharmacological activity of novel h-E3
 CC UPL via administration of a modulator or heterologous expression, is used
 CC for treating physiological conditions like inflammation, autoimmune
 CC diseases, neurological disease, apoptosis, endothelial cell physiology
 CC (e.g., proliferation, differentiation), peripheral vascular disease,
 CC angiogenesis, cancer, hematopoietic disorders, arthritis, cachexia,
 CC leukemia, pulmonary disorders, diabetes and viral infection. The
 CC nucleotide sequences which encode h-E3 UPL may also be employed in
 CC analysis to map chromosomal location e.g., screening for functional
 CC association with disease markers. They are also used as screening tools
 CC in the identification of appropriate human subjects and patients for

therapeutic clinical trials. The sequences can also be used to detect the presence of the mRNA transcripts in a patient or to monitor the modulation of transcripts during treatment. This sequence represents the human E3 ubiquitin protein ligase protein described in the method of the invention.

Sequence 852 AA:

Query Match 33.5%; Score 1301; DB 20; Length 852;
Best Local Similarity 34.7%; Pred. No. 3.2e-118;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

6 KRLVYLVAK-NLAKDPFRLDPFPAKIVVDGSGCHSTDTVKTLPDKNQHYDLYVGK 64
9 QLOITVISAALKENKKNKMGF-PSPYVEVVD--GQSKTEKKNNTNSPKMQPLVITVP 65
65 TDSITISVNNHKKHK--QGAGPFCVRLSNASRLKDT-----GYGK-----L 108
66 VSKLHFRWVSHQTLKSDVLLGTAALDIYETLKNMKLEVVVTLQLGDKPEPTETIGDL 125
109 DLCKLPSPDTAVRQIVVSLQT-----RDRIGTGS-----VYDC 144
126 SIC---LDGLQLESIVTNGETTESASQNDGSRKMDETRVSTNGSDDPDAGAGEN 181
145 RGL-----LENEGTVEDSGRPLSCFMEEPAPYDSTGAAGGNCFFVS----- 192
182 RRVSGNNSPBLSSNG--FKPSRPRP--SRPPPTPRRPAVNSGSPSATSSDSST 234
193 -----PSCDRLQAQRLNPD 208
235 GSLPPTNTNTSEAGTSGLIPLTISGSGSPRLNPVYQAPLPQMGQRV-----D 286
209 VRSSL-----QTQNRPHGHQSPLEPEGYEORTVQGVYFHTGTYSTHNDPRIP 260
287 QHGRVYVYDVHEKRTWDRE--PLPPGWRVNDNGRYYVDHPTKRTTQORPLLE 341
261 -----RDL-----NSVNCDELGPFPQMEVSTVSG 286
342 SVANVEGMQORQOLGAMQGFQRIYGNQDLFATSQSKEDPLGRLPPMEKRTDSNG 401
287 RIFVVDHNNFTTFTDRLHIMNHOCQKEPSQPLPL----- 324
402 RYFVNNHTRIITWEDR-----SQGLNE--KPLPEGEMKRTVDSIPYVDHNRRTT 453
325 -----PSEGLSEDEELPAQRYERDVOKLVLR--HELSLOQPAQGHORIEVSREIPER 377
454 TYIDPRGKALNGQIAYVDFRKNQVFRKQQLAMPQ---HKTIVTKRTLPED 509
378 SYRQIMKRRPKDKKLMVYKRGESGLDYGVAREWLYLICHEMLNPPYGLFQYSTDNLY 437
510 SFOQIMSFSPQDLRRRLWVTFPEEGGLDYGVAREWFFLSHEVINPMVCLFEYAGKDNV 569
438 MLOINPDSINPDLSEFHFVGRIMGLAVFHGYINGSTVFYKQLLQKPLQSDSESV 497
570 CLQINPASVYNPHLKYFRIFGFMALPHGKFLDTGSLSPFYKRIANKPYGLKDLDSI 629
498 DPELHSLVWILENDITPV-LDHTPCVENHARGLLOHELEKNGNVPYTEENKKEYVRL 556
630 DPERYNLIWKENNIEBCLDEWYFSDVKEIIGELIKSHDLKNGNINIVTEENKEYIRM 689
557 YAMRFRGIEAGFALQKGFNELLPOHLKPPDOKEILLIGLDKIDINDMKNSTELK 616
690 VAEWRUSRGVEEQTQAFBGFENILPQOYLQFDALKELEVLQGMQEDLNDWQHAIYR 749
617 HCVADSNIVFMQAVETPDEERRARLLQFVGTSTVPQLQGFALQSGTGAAGPRLFTIH 676
750 HYARTSKQIWMFQVKEIDNEKRMKLLQFVGTCLPLVPGFADLMGNSN---GPKKFCIE 806
677 LIDANTNLPKATCPRRIDIPYSEYELVETLLTAVETGGFAVE 723
807 KV-GKENWLPFRSHCFNRDLDPYKSYEOLKEKLLFALETETGFOE 852

RESULT 15

AA25170
ID AA25170 standard; Protein: 927 AA.

AA25170;

09-SBP-1999 (first entry)

Human KIAAN ligase protein fragment.

Inhibitor: ubiquitin-mediated proteolysis; IkappaB polypeptide; HECT; ubiquitin-conjugating system; Homologous to Bc-Ap carboxyl terminus; ligase; ubiquitin; ubiquitination; RSC ligase; KIAAN ligase; UbC4; MAD3; modulator; treatment; proliferative disorder; apoptosis; sepsis; differential disorder; viral infection; tissue wasting disorder; cachexia; malignancy; inflammatory disease; parasitic disease; tuberculosis; IL-2 therapy; rheumatoid arthritis; rheumatoid spondylitis; osteoarthritis; gouty arthritis; respiratory distress syndrome; cerebral malaria; chronic pulmonary inflammatory disease; sickle cell; pulmonary sarcoidosis; bone resorption disease; reperfusion injury; graft versus host reaction; allograft rejection; Crohn's disease; ulcerative colitis; pyresis; multiple sclerosis; autoimmune diabetes; systemic lupus erythematosus; leprosy; AIDS.

Homo sapiens.

Key Location/Qualifiers
FT Protein 1..927
/note="partial sequence, no start codon given"

MO9904033-A1.

28-JUN-1999.

16-JUL-1999; 98MO-US14638.

16-JUL-1997; 97US-0895601.

(MITO-) MITOTIX INC.

Beer-Romero P, Glaes SJ, Rolfe M, Strack PR;

WPI; 1999-132274/11.

N-PSDB; AAX78494.

Identifying modulators of I-kappa-B proteolysis - used to develop products for treating e.g. proliferative and/or differentiative disorders, infections, tissue wasting, cachexia or AIDS

disclosure; Page 68-71; 79pp; English.

This invention describes a novel assay identifying an inhibitor of ubiquitin-mediated proteolysis of a IkappaB polypeptide. The method comprises: (a) providing a ubiquitin-conjugating system including the IkappaB polypeptide and a HECT (homologous to Bc-Ap Carboxyl Terminus) ligase and ubiquitin, to promote ubiquitination of the IkappaB polypeptide by the HECT ligase; (b) contacting a level of ubiquitination of the polypeptide in the presence of the candidate agent and (c) comparing the measured level of ubiquitination in the presence of the candidate agent with ubiquitination of the IkappaB polypeptide in the absence of the candidate agent; where a statistically significant decrease in ubiquitination of the IkappaB polypeptide in the presence of the candidate agent is indicative of an inhibitor of ubiquitination of the IkappaB polypeptide. The assays can be used to identify compounds which modulate binding and/or ubiquitination of an IkappaB (or other cellular or viral substrate) by a HECT ligase, such as RSC or KIAAN. Such modulators can be used e.g. in the treatment of proliferative and/or differentiative disorders, to modulate apoptosis in the treatment of viral infections, and in the treatment of tissue wasting disorders e.g. cachexia secondary to infection or malignancy, cachexia secondary to human AIDS, inflammatory diseases, parasitic diseases, tuberculosis and high dose IL-2 therapy; rheumatoid arthritis, rheumatoid spondylitis,

CC osteoarthritis, gouty arthritis and other arthritic conditions, sepsis,
CC respiratory distress syndrome, cerebral malaria, chronic pulmonary
CC inflammatory disease, silirosis, pulmonary sarcoidosis, bone resorption
CC diseases, reperfusion injury, graft versus host reaction, allograft
CC rejections, Crohn's disease, ulcerative colitis, or pyresis, in addition
CC to a number of autoimmune diseases such as multiple sclerosis, autoimmune
CC diabetes, systemic lupus erythematosus; and ENL in leprosy, HIV, and
CC AIDS. This sequence represents a human KIAAN ligase which is used in the
CC method of the invention.
XX
SQ

Sequence 927 AA:

Query Match 32.4%; Score 1259; DB 20; Length 927;

Best Local Similarity 33.6%; Pred. No. 5.1e-114;
Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

QY 7 IRLTYCAKMLAKKDPFRLLDPPEAKIV---VDSSGQCHSTDTVNTLDPKKN----- 55
DB 48 VRVAVIAGIGLAKKDIIGASDPYVTVLYDPMGVLTISVQTKIKSLNPKMNEILFRV 107
QY 56 ---QH-----YDLVYKTD--SITISVWN-----HKIKKKOG 83
DB 108 HPQOHRLLPEVPEBNELTRDDPLGQVDPVLYPLPTENPLERPYTFKDFVLHPSHSKRV 167
QY 84 AGFLGCVRLISNAISRLKDT-----GYQLD---LCKL----- 113
DB 168 KGYL-----RLKKTLYLPKTSGEDNAEQAELEPEGMVVLDPDPACHLQOQOE 216
QY 114 -----NPSD--TDVARGIVSLQ---TRDR 134
DB 217 PSPLPPEWEERQDILGRITYVNHESRRTQMKRPTPODNLTDENGNLQIQORAFTRRQ 276
QY 135 IGTGGSVDCRGLLEN-----EGTVYEDSG--PGRPLSCFMEBPAYTD----- 176
DB 277 ISEETESVDNQESSEWEIIRDEEATMYSSQAFSPSPSSNLDVPTHLAEHLNARLTIFG 336
QY 177 ----STGAAGGNCR-----FVESPS-----QDQRLQAQR-- 204
DB 337 NSAVSQPASSNSHSSRGSLOAYTFEEQPTLPVLLPTSSGLDPGWEKQDERGSRSYVDH 396
QY 205 -----RNPVNGSLQTPQ-----NRPHGOSPELPEGYEQ 234
DB 397 NSRTTWTKPTVATVETISQLTSSQSSAGPQSQASTSDGQVQTQPSLEHGFPEKWEV 456
QY 235 RTTVQGVVFLHTQGVSTHDP--IPRDLNSV---NCDELGPLPGWEVSTVSQRI 288
DB 457 RHAENGRPFIDNITKTITWEDPRKIPALHGRKTSLDTSNDLGPLPGWEERTHTDRI 516
QY 289 YFVDHNNRTQFTDPRLHIMNHQCOLKEPSOPLPLPSEGSLEDEELPAQRYERDLVOKL 348
DB 517 FYTNHNKIKTQWEDPRLENV-----AITGPAYP-----YSRDYKRY 553
QY 349 KVLRLHELSLQOPAGHCRLEVSREEIFESSYROIKMRPKD--LKRLMVKFRGEGLDYG 407
DB 554 EFERRLKKQNDIPNKEEMKLRRATVLEDSYRIMGVKRADELKARLWIEFDEGEKLDYG 613
QY 408 GVAHEMLYLCHMLNLYYGLFOYSTDNLYMLOINPDSSI--NPDLSYHHFVGRIMGLAV 466
DB 614 GVAHEMFLLSKENENFYGLFEYSATDNTYLLQINPNSGICNEHLSYFKFGRVAGAV 673
QY 467 FGHYINGGFTVPFYKQLGKPIQLSDLSVDELKSLWILENDITPVLDHTEPCVENH 526
DB 674 YHGKLDGPFIRPFYKMLHKPITLHDMESVDSEYNSLRWILENDPTE--LDLRFIIDE 732
QY 527 AFGRILOHELKPRGNRPVTEENKKEYVLLYVWRPMRGIEAOFLALOKGFNELIPOHL 586
DB 733 LFGQTHQHELKNGGSEIVTNKKNKEIYIVIQMRVNRIOKMAAFKGFPELLPODI 792
QY 587 KPPDQKELEIIGGLDKIDLNDKSNTRLNG--CVADSNIVRWFWQAVETFDERRARLLQ 645
DB 793 KIPDENLEILMCLGADVVDWREHTKYNGYSANHOVIOWFWKAVLMDSEKRIILLQ 852
QY 646 FVTGSTRVPLQGFRAAGSTGAAGPRFLTHILDANTDNLKRAHTCFNRIDIPYESTEK 705

DB 853 FYTGSRVPMNGFAELIYGSN---GPGSFVQW--GTPEKLPRAHTCFNRIDLPPYSFEE 908
QY 706 LYEKLLTAVESTCGF 720
DB 909 LMDKLIQMAIENTQGF 923

Search completed: February 20, 2004, 15:28:29
Job time : 44.3032 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 15:26:11 / Search time 15.2366 Seconds
(without alignments)
2007.718 Million cell updates/sec

Title: US-10-009-945-2
Perfect score: 3884
Sequence: 1 GGSSIKIRLVLCARLAKK.....ELTYEKLTAVETGFAVE 723

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 4231058 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2913.5	75.0	735	3	US-08-539-205A-2
2	2913.5	75.0	735	4	US-08-392-163A-2
3	1471	37.9	766	3	US-08-539-205A-4
4	1471	37.9	766	4	US-08-392-163A-4
5	1318	33.9	854	2	US-08-070-060-4
6	1318	33.9	854	3	US-08-357-746-4
7	1301	33.5	852	2	US-09-070-060-3
8	1301	33.5	852	3	US-09-357-746-3
9	1259	32.4	927	3	US-08-895-601-6
10	1228.5	31.6	834	4	US-08-539-205A-6
11	1228.5	31.6	834	4	US-08-392-163A-6
12	1225	31.5	906	3	US-08-630-916A-48
13	1047	27.0	683	3	US-08-630-916A-46
14	530.5	13.7	874	2	US-08-247-904B-8
15	530.5	13.7	874	3	US-08-767-942A-21
16	509	13.1	866	2	US-08-100-692-1
17	509	13.1	866	2	US-08-674-030-1
18	473	12.2	1083	3	US-08-895-601-5
19	241	6.2	486	3	US-08-348-518C-2
20	238.5	6.1	472	3	US-08-348-518C-5
21	238.5	6.1	472	3	US-08-476-509B-5
22	162.5	4.2	54	3	US-08-630-916A-124
23	162	4.2	448	3	US-08-476-509B-2
24	161.5	4.2	454	3	US-08-348-518C-4
25	154.5	4.0	454	3	US-08-476-509B-4
26	154.5	4.0	55	3	US-08-476-509B-75
27	148.5	3.8	58	3	US-08-630-916A-84

28	148	3.8	51	3	US-08-630-916A-117	Sequence 117, App
29	147.5	3.8	54	3	US-08-630-916A-74	Sequence 74, App
30	145	3.7	51	3	US-08-630-916A-73	Sequence 73, App
31	141	3.6	38	3	US-08-630-916A-36	Sequence 36, App
32	140	3.6	54	3	US-08-630-916A-118	Sequence 118, App
33	139	3.6	51	3	US-08-630-916A-116	Sequence 116, App
34	138	3.6	1105	4	US-08-710-249-2	Sequence 2, App
35	138	3.6	1105	4	US-09-220-157A-2	Sequence 2, App
36	133	3.4	224	3	US-08-630-916A-50	Sequence 50, App
37	132	3.4	51	3	US-08-630-916A-115	Sequence 115, App
38	130	3.3	38	3	US-08-630-916A-24	Sequence 24, App
39	130	3.3	38	3	US-08-630-916A-26	Sequence 26, App
40	130	3.3	38	3	US-08-348-518C-15	Sequence 15, App
41	130	3.3	38	3	US-08-348-518C-18	Sequence 18, App
42	130	3.3	38	3	US-08-476-509B-15	Sequence 15, App
43	130	3.3	38	3	US-08-476-509B-18	Sequence 18, App
44	130	3.3	335	2	US-08-844-312-2	Sequence 2, App
45	128	3.3	38	3	US-08-630-916A-25	Sequence 25, App

ALIGNMENTS

```
RESULT 1
US-08-539-205A-2
/ Sequence 2, Application US/08539205A
/ Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nelsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-539-205A-2
Query Match 75.0%; Score 2913.5; DB 3; Length 735;
Best Local Similarity 72.9%; Pred. No. 1.4e-280;
Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;
```

Db 59 SDSVTISVNMHKKIKHKKQAGFLGCVRLLSNAINRLKDTGYQRDLCLGPNNDTVRGQ 118
 QY 125 IVVSLQTRDRIGTGSSVYDCRGLLENE-----GTVY----- 155
 Db 119 IVVSLQSRDRIGTGSSVYDCRGLLENE-----GTVY----- 178
 QY 156 --EDSGRPLSCFMEBPAYTSTGAAGGNCRFVESPQDQRLQARLNPVRSGL 213
 Db 179 ASFYSSGRLSCFVDENPISGTNATCG-----QSDPRLARRRARSQHRHYM 229
 QY 214 QTPQNRPHQSPLEBEGYQRTTVQGVYFLHTQVSTWHPRIPLDLSVNCDELGP 273
 Db 230 ---SRTHLHTPDLPEGYEQRTTQGVYFLHTQVSTWHPRIPLDLSVNCDELGP 285
 QY 274 LPPGWEVRSVSGRIYFVDDNNRTTQFTDPR---LHIMNQCLKEPSQPLPSEGS 329
 Db 286 LPPGWEIRNTATGRVYFVDDNNRTTQFTDPRLSANLHLVLRNQKQOQOQVY---S 341
 QY 330 L---EDELPAQRYERDLYQKLVLRHLSLQOPQAGHCRIEVSREIFESYQIMQR 386
 Db 342 LCPDTECLTPRYKRDLYQKLVLRHLSLQOPQAGHCRIEVSREIFESYQIMQR 401
 QY 387 PKDLKRLMKVFRGEGLDYGVAREMLYLCHENLNPYGLFOYSTNIYMLQINPDS 446
 Db 402 PKDLKRLMKVFRGEGLDYGVAREMLYLCHENLNPYGLFOYSTNIYMLQINPDS 461
 QY 447 INPDLSYFHFVGRINGLAVERHGYINGFTVBFYKQLGKPIQSDLESVDPPELHSLV 506
 Db 462 VNPBLSYFHFVGRINGLAVERHGYINGFTVBFYKQLGKPIQSDLESVDPPELHSLV 521
 QY 507 WILNDITPVLDHFCVEHNAFRLLOHEIKPGRVNPVTEENKKEYVRLYNNRFPNGI 566
 Db 522 WILNDITGVLDHFCVEHNAFRLLOHEIKPGRVNPVTEENKKEYVRLYNNRFPNGI 581
 QY 567 EAQFLAQKGFNEILPOHLKPEDOKELELIGLKDIDLMKNTRLKHCVAASNIVR 626
 Db 582 EAQFLAQKGFNEILPOHLKPEDOKELELIGLKDIDLMKNTRLKHCVAASNIVR 641
 QY 627 WFOQAVETFEERRARLLQVYTGSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 686
 Db 642 WFOQAVETFEERRARLLQVYTGSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 698
 QY 687 KAHTCFNRIDIPYSEYEKLYEKLTAIVETCGFAVE 723
 Db 699 KAHTCFNRIDIPYSEYEKLYEKLTAIVETCGFAVE 735
 RESULT 2
 US-09-392-163A-2
 ; Sequence 2, Application US/09392163A
 ; Patent No. 6503742
 ; GENERAL INFORMATION:
 ; APPLICANT: Beach, David H.
 ; APPLICANT: Caligiuri, Maureen
 ; APPLICANT: Nelesky, Bradley
 ; TITLE OF INVENTION: Ubiquitin ligases, and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/392,163A
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: CSV-005.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 735 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-392-163A-2
 Query Match 75.0%; Score 2913.5; DB 4; Length 735;
 Best Local Similarity 72.9%; Pred. No. 1,48-280;
 Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;
 QY 5 IKRLTVLCAKMLAKKDFRLLPDPFAKIVDGSQCHSTDTVKNLTPKNNQHYDLYGK 64
 Db 12 VKRLTL-----GLPDPFAKIVDGSQCHSTDTVKNLTPKNNQHYDLYGK 56
 QY 65 TDSITTSVNMHKKIKHKKQAGFLGCVRLLSNAINRLKDTGYQRDLCLGPNNDTVRGQ 124
 Db 59 SDSVTISVNMHKKIKHKKQAGFLGCVRLLSNAINRLKDTGYQRDLCLGPNNDTVRGQ 118
 QY 125 IVVSLQTRDRIGTGSSVYDCRGLLENE-----GTVY----- 155
 Db 119 IVVSLQSRDRIGTGSSVYDCRGLLENE-----GTVY----- 178
 QY 156 --EDSGRPLSCFMEBPAYTSTGAAGGNCRFVESPQDQRLQARLNPVRSGL 213
 Db 179 ASFYSSGRLSCFVDENPISGTNATCG-----QSDPRLARRRARSQHRHYM 229
 QY 214 QTPQNRPHQSPLEBEGYQRTTVQGVYFLHTQVSTWHPRIPLDLSVNCDELGP 273
 Db 230 ---SRTHLHTPDLPEGYEQRTTQGVYFLHTQVSTWHPRIPLDLSVNCDELGP 285
 QY 274 LPPGWEVRSVSGRIYFVDDNNRTTQFTDPR---LHIMNQCLKEPSQPLPSEGS 329
 Db 286 LPPGWEIRNTATGRVYFVDDNNRTTQFTDPRLSANLHLVLRNQKQOQOQVY---S 341
 QY 330 L---EDELPAQRYERDLYQKLVLRHLSLQOPQAGHCRIEVSREIFESYQIMQR 386
 Db 342 LCPDTECLTPRYKRDLYQKLVLRHLSLQOPQAGHCRIEVSREIFESYQIMQR 401
 QY 387 PKDLKRLMKVFRGEGLDYGVAREMLYLCHENLNPYGLFOYSTNIYMLQINPDS 446
 Db 402 PKDLKRLMKVFRGEGLDYGVAREMLYLCHENLNPYGLFOYSTNIYMLQINPDS 461
 QY 447 INPDLSYFHFVGRINGLAVERHGYINGFTVBFYKQLGKPIQSDLESVDPPELHSLV 506
 Db 462 VNPBLSYFHFVGRINGLAVERHGYINGFTVBFYKQLGKPIQSDLESVDPPELHSLV 521
 QY 507 WILNDITPVLDHFCVEHNAFRLLOHEIKPGRVNPVTEENKKEYVRLYNNRFPNGI 566
 Db 522 WILNDITGVLDHFCVEHNAFRLLOHEIKPGRVNPVTEENKKEYVRLYNNRFPNGI 581
 QY 567 EAQFLAQKGFNEILPOHLKPEDOKELELIGLKDIDLMKNTRLKHCVAASNIVR 626
 Db 582 EAQFLAQKGFNEILPOHLKPEDOKELELIGLKDIDLMKNTRLKHCVAASNIVR 641
 QY 627 WFOQAVETFEERRARLLQVYTGSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 686
 Db 642 WFOQAVETFEERRARLLQVYTGSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 698
 QY 687 KAHTCFNRIDIPYSEYEKLYEKLTAIVETCGFAVE 723
 Db 699 KAHTCFNRIDIPYSEYEKLYEKLTAIVETCGFAVE 735

RESULT 3
US-08-539-205A-4
Sequence 4, Application US/08539205A
Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nefsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-7000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-539-205A-4

Query Match 37.9%; Score 1471; DB 3; Length 766;
Best Local Similarity 40.9%; Pred. No. 5.4e-137;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

QY 4 SIKIRLTVLCARLAKKDPFRLPDPRAKIVVDGSGCHSTDTVKTLPKRNQHDLYG 63
DB 7 SRIRIVIVADLYKRDVFRFPDPVALTVGE-QTHTTAKIKTLNPMWNETFEVAVT 65
QY 64 KTDSTITSVNHHKIKKQAGFLGCVRLLSNAISRLKDTGYQL--DLCKLNPSTDAV 121
DB 66 DNSTIAIQVFDQKFF-KKKQGGFLGIVINRQGVLDLAIQGDGEMTRDLKSN--ENTVV 122
QY 122 RQGIIVSLQ-----TRPRITGSSVYDCGLNESTVED 157
DB 123 HGKIILNSTAOSTQVPSAASGARTORTSITNDPQSSKSSVSRNPASSRASPFRD 182
QY 158 SGPG-----RPLSCFME--PAY--TDSG-----AAA 182
DB 183 NAPAASPASSSEPTFSEFDQGRLLPGEWERTDNGRTYYVDNHTRTTWRPLSSVA 242
QY 183 GGNCRFVESPS-----QDQRLAQSLRNDVVGSLQTPQNRPHGQSPLEEGTEQ 234
DB 243 GAAAEHLSSASSANVTEGVQPSSSNAARTEASVLTSNATAG-----SGELPPGMEQ 296
QY 235 RTTVQGVFLHTQTGVTWHDPRIRPDINSVY-----CDELGPLPGEVEST 283
DB 297 RYFEGRPYFVDNHTKRTTVDPRRQGYIRSYPNNATIQQAPVSLGRLPSGEMRLT 356
QY 284 VSGRIYFVDNHTRTQFTDRLHHNMHQCCKLEPSQPLPSEGLSDDELPAQRYRD 343
DB 357 NTAAYVEVDNHTKRTTWDPR-----LPS--SL--DQNVF--QYKRD 392

QY 344 LVQKTLVRLHELSLQOPQACHRIEVSRESEIFSESYROIKMKRPKDLKKRLMYKFRGEEG 403
DB 393 FRKLIYPLSGPAL-HPLPGQCHKVTRRNHIFEDSVAINRQASTDIAKRLIMKIFDSEEG 451
QY 404 LDYGVAAEYMLYLCHEMLNPDYGLFOYSTDNITMLQINDPSSINPDHLSFFHVGRIKG 463
DB 452 LDYGLSREYFELLSHENFPYCLFEYSVDNTTLQINPHSGINPEHLVYFPIGVIG 511
QY 464 LAVPHGYNIGFTVPEFYKOLGKPIQSDLESYDPPELHKSILWILENDITPYLDHTFCV 523
DB 512 LAIFHRFPDAPFVVSFIMKILQKVTIQMESDADAYISLWILDNDITGVLDLTFV 571
QY 524 ENNAFGRILOHELRKNGNVDPTEENKKEYRLVYNNRPMRGIEAQLAQGPNELIPQ 583
DB 572 ENCGEVVITIDLKNGRNIEVTENKREYDVLTVW-IQRIEEQNAFHEGFSSELIPQ 630
QY 584 HLKPEPDKELELIIGLDKTIDLDNWSNTRLKHCVADSNIVRMFOAVFPEBRRL 643
DB 631 ELINVPDERELELIGISEIDMEDWKHKDYRSENDQIIFKFMLEWBSNKKSR 690
QY 644 LQFTVGTSTVPLQGFKAQSGTGAAPRLFTIHLIDANTDNLPRAHCFNRIDIPYEST 703
DB 691 LQFTGTSTRIPVNGFKDQSD--GPRKFTIEKA-GEPMKLPRATCFNRDLDPYTSK 746
QY 704 ELVYEKLTAVEETGFAVE 723
DB 747 KDLHKLSTAVEETIGFQGE 766

RESULT 4
US-09-392-163A-4
Sequence 4, Application US/09392163A
Patent No. 6503742
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nefsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163A
FILING DATE:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/539,205
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-392-163A-4

Query Match 37.9%; Score 1471; DB 4; Length 766;

Best Local Similarity 40.9%; Pred. No. 5.4e-137;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

```

QY 4 SIKIRLTVLCANLAKOFFRLPPAKIVVDGSGQCHSTDTVTKNTLDPKMNQHYDLYG 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 SRIRVTVIADGLYKRDVFRFPDPFAVLTVDS- QTHTAIKKTLNFWNNEFEVNTV 65
QY 64 KTDSTISVMNHKKIKKQAGAGLGVLLSNLSRLKDTGYORL- DLCKLNSDPAV 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 DNSTAIQVFDQKF- KKKGGFLVYNLRGVDDVLALGDEMLTRDLKSN- EMTTV 122
QY 122 RQGIIVSLQ-----TRDRIGTGGSVVDRGLNEGTVEYD 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 HGKILINLSTASTLOVPSASAGARTQRTSITNDPOSSKSSVSBNPASRAGSPTRD 182
QY 158 SGPG-----RPLSCFME- PARY- TDSGT-----AAA 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 NAPAASPASSEPRTSSFEQYGRLPFGHERITDMLGRTTYVDHNTASTTWIRPNLSVA 242
QY 183 GGCNCFVESPS-----QDRLQARLNDPDVRSGLQTPONRPHGSGPELFEQYEQ 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 GAAPALHSSASANTVEGVQSSSMAARTEASVLTSNATTAG-----SGELPQMEQ 296
QY 235 RTVQGVQVFLHQGVSTWMDRIPLDLSVN-----CDELGRPLPGMEVRS 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 RYTPGRPFVVDHNRRTTWVDRQYRSTYGGNNATIQQPVSGPLSGMEWELT 356
QY 284 VSGRIYFVDHNRRTQFTDRLHIMNQCLKEPSQPLPSPSGSLEDELPAQRYERD 343
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 NTAHYVFDHNTTTWDDPR-----LPS--SL-DQVVP--QYKRD 392
QY 344 LVOKLKVLEHLSLOQPOAGCRIEVSEIEFEESYRQIMKRPDLKRLMVFREBEG 403
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 393 FRRLIYFLSQPL-HPLGQCHIKVRNHLFEDSYAIMEQSATDLKRLMIFDEBDG 451
QY 404 LDYGVAREMLYLICHEMLNPPYGLFOYSTDNIVMLQINPSSINPDLSYFHFVGRMG 463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 452 LDYGLSREYFLLSHENMFYCLFEYSVUNYLLQINPHSGINPELNFKIGVIG 511
QY 464 LAVFHGYINGGTVPFYKQLISDLDSVDEPLHKLWLINDITFPVLDHFFCV 523
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 512 LAIFHRFPVDAFVVSFYKQILQCKVTLQDMESMDAEYRSYLVMLDNDIGVLDLFFSV 571
QY 524 EHNAFRIILOHEIKPRGRVVPTEENKERYRLYNNRPMGIEAQFLAQGFNEELIPQ 583
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 572 EDNCFGEVTTIDKPRGRNIEVTEENKERYVDLVVW-IQKRIEQFNAFEGSELIPO 630
QY 584 HLKPFQKELIILIGLDKIDLNMKSNTRLKHCVADSNIVRMFQAVETFEDEERARL 643
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 631 ELINVFDERELLELLIGISEIMEDMKKHDYRSSENDQIILKMFWEIMDEWSNEKXRL 690
QY 644 LQFVGTSTVPLQGFALQSGTAGAPRLFTIHLDAINTDLPKANTGENIDIPVRESY 703
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 691 LQFTTISRIPVNGFDLQSD--GPKFTLEKA-GEENKLPKRAHTOFNLDLPTTSK 746
QY 704 EKLXKLLTAVERTCGFAVE 723
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 747 KDLDKLSTIAVEETIGQGE 766

```

RESULT 5
US-09-070-060-4
Sequence 4, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Husbard, Carolyn M.
TITLE OF INVENTION: Human B3 Ubiquitin Protein
TITLE OF INVENTION: Ligase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENBA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike

```

? CITY: Wilmington
? STATE: DE
? COUNTRY: USA
? ZIP: 19850-5437
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/070,060
? FILING DATE: 30-APR-1998
? CLASSIFICATION:
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 60/073,839
? FILING DATE: 05-FEB-1998
? ATTORNEY/AGENT INFORMATION:
? NAME: Higgins, Patrick H
? REGISTRATION NUMBER: 39,709
? REFERENCE/DOCKET NUMBER: PHM, 70312
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 302.886.4889
? TELEFAX: 302.886.8221
? TELEX:
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 854 amino acids
? TYPE: amino acid
? STRANDEDNESS: unknown
? TOPOLOGY: unknown
? MOLECULE TYPE: peptide
? US-09-070-060-4

Query Match 33.9%; Score 1318; DB 2; Length 854;
Best Local Similarity 35.4%; Pred. No. 1.1e-121;
Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

QY 6 KIRLVLCAL-NLAKDPFRLPPFAIVVDGSGQCHSTDTVTKNTLDPKMNQHYDLYGK 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 QIQTIVIAKAKENKKKWFQ-PSRYVEVTD--QGSIKTKCNNTSPKKQPLATVIVP 65
QY 65 TDSITISVMNHKKIKKQAGAGLGVLLSNLSRLKDTGYOR-----107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 TSKLCFRVMSQTL--KSDVLTAGLDIYETLKSNNMLLEVWNTLQVGRKEPTEEM 122
QY 108 --LDLCKLNPSTDAVAGQIVYS-----LQTRD--RIGTGS--VVDGR 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 GDLVYC--LDGLQVEAEVVTNGETSCSESTTQNDGCTRTDTRVSTNGSEDEPEVAAS 178
QY 146 G-----LLENQTVYEDSGPRPLSCFMEBPAYT-----DSTG 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 GENKRNAGNNSPSLSNG--FKPSRPFR-----SRPFPTRPRPASVNGSPSTNSDSG 231
QY 180 AA-----AGGNRPFVESQ-----DQRLQAR 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 SSTGSLPPTNTNNTSTSEGTSGLLIPLTISGSGRPLNTVSQLPFGMGQRY-----287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 LKMPVRSGL-----QTFONRPHGSGPELFEQYGRFTTVQGVYPLATQGVSTW 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 ---DQHRVYVVDHVEKRTWDRP-----PLPFGWERVVDNMGRRYYVDHPTRTTQQ 338
QY 256 DFRTP-----RDV-----NSVNCDELGRPLPGMEVR 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 RPLESVANVEQWOLQROSLQAGAQCFNQGFIYGNODLFATSONKEFDPLGLPFGMEKR 398
QY 282 STVSGRIYFVDHNRRTQFTDRLHIMNQCLKEPSQPLP-----324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 TDSNCRVYFVNHNRRTQMEDPR-----SQGLNE--KPLPFGMEKRFTVQGIYFVDH 450
QY 325 -----PESGSLDEBELPAQRYERDLYOKLKYLA--HELSLOQPOAGCRIEVSE 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 451 NRRATTVIDPRTGSAALDNGPOLAYVADFRAKVQYFFRWQQLAMPO--HIKITVTRK 506

```

QY 373 EIPESYQIMKRPKDLKRLMVKFRGEGLDYGVAREMLYLICHEMLNPFYGLFOYS 432
DB 507 TLFEDSFQOIMSPBQDLRLRLMVFIPGEGLDYGVAREMFLLSHETLNMVCLFEYA 566
QY 433 TDNIYMLQINPDSSINDHLSTFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLGKPIOLS 492
DB 567 GKDNVYCIQINPDASYINPDHLKYFRFGRFAMALPHGKFDIGFSLPFYKRIINLPVGLK 626
QY 493 DLESVDEBELKSLVWILENDITPV-LDHTFCVHNAPFGRILQHEIKPNRANVPYENKK 551
DB 627 DLESIDPEFINSILWVKNENIEBCEGLMVFSDXKILGEIKSHDLKPNGNILVTEENKE 686
QY 552 EYRLVYNNRFRMGIEAQLALOKGFNELIPQHLKRPDQKLELIIIGLDKIDINDMKS 611
DB 687 EYRMVAMERLSRGVEEQTAFFEGFNEILPQOYLQYFAKLELVLLCGMEIDINDMQR 746
QY 612 NTRLKHCVAADSNIVRMQAVETFEDEERARLIQFTMGSTRVPLQGFALQOSTAGAPR 671
DB 747 HAIYRHYTRTSKQIMFWQFVKEIDNEKMRLLQFTVGTGRLPVGGFADLMGNSN--GPQ 803
QY 672 LFTIHLIDANTDNLPKAHTCFNRIDIPYSEYEKLYEKLTAVEETCGFAVE 723
DB 804 KFCIEKV-GKENWLPKSHTCFNRIDLPYKSYQKLEKLLFAIEETEGFGE 854

RESULT 6

US-09-357-746-4
Sequence 4, Application US/09357746
Patent No. 6087122
GENERAL INFORMATION:
APPLICANT: ZENCA Limited
TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
FILE REFERENCE: PHM.70312.N1
CURRENT FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: US No. 6087122 60/073, 839
EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: US No. 608712209/070,060
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 4
LENGTH: 854
TYPE: PRT
ORGANISM: Mus musculus
US-09-357-746-4

Query Match 33.9%; Score 1318; DB 3; Length 854;
Best Local Similarity 35.4%; Pred. No. 1,le-121;
Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

QY 6 KIRLVTLCAK-NLAKQDFRLLDPPAKIVYDGGQCHSTDTVKNLTDPKKNCHYDIYVK 64
DB 9 QLOITVISAUKLKKENKKNFQ--PSPIVEYTVD--GQSKTKKCNNTNSPKKQPLVIVTP 65
QY 65 TDSITISYWNHKKIKHKKAGAGFLGCVRL-----LSNAISRLKDTGYR----- 107
DB 66 TSKLCFRVWSHQTL--KSDVLLGTAGADIVETLKSNNMKLEEVVMTLQVGDKEPLETM 122
QY 108 --LDLCKLNPSTDAVRGQIVS-----LQTRD--RIGTGG--VYDGR 145
DB 123 GDLVVC---LDGLQVEAEVAVTNGEETSGESTGNDGCRDTRVSTNGSDPEVAAS 178
QY 146 G-----LLENEGTVEDSGPGRPLSCFMEBPAPY-----DSTG 179
DB 179 GENKRANGNNSPSLNG--FKPSRPFR-----SRPPPTPRRPAVNGSPSTNSDSG 231
QY 180 AA-----AGGNCRFVSEPG-----DQRLQAR 203
DB 232 SSTGSLPPTNNTVNTSTEGATGILPLTISGSGRPPLNTVSGQAPLPGMEQFV---- 287
QY 204 LNPVVRGSL-----QTFQNRPHGHQSPELPEGEQRTTVGQVYFHTOTGVSTWH 255

DB 288 ----DQGRVYVYDVHVKRTTWDRB-----PLPGRWERVDNMGRIYYVDHFTRTTWQ 338
QY 256 DRIIP-----RDJ-----NSGNCDELGLPGRGVR 281
DB 339 RPTLESVRNVEQMLQORSQLQAMQOFNORFIYGNQDLPATSONKEPPLGSLPGRGWEKR 398
QY 282 STVSGRIYEVVDHNNRTQPTDRLHHIMNHQCOLKEPSCQPL----- 324
DB 399 TDSNGAVYVFNHTRTQEDR-----SQQLNE--KPLPEGEMKFTVDGIPYVDH 450
QY 325 -----PSESLEDELPRQREPRDLVQKLYR--HELISQQAHCNIEVRE 372
DB 451 NRDATYIDPRGKSLDNGPQIAYVRPFAVQYFRWCQQLAMPQ--HIKITVRK 506
QY 373 EIPESYQIMKRPKDLKRLMVKFRGEGLDYGVAREMLYLICHEMLNPFYGLFOYS 432
DB 507 TLFEDSFQOIMSPBQDLRLRLMVFIPGEGLDYGVAREMFLLSHETLNMVCLFEYA 566
QY 433 TDNIYMLQINPDSSINDHLSTFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLGKPIOLS 492
DB 567 GKDNVYCIQINPDASYINPDHLKYFRFGRFAMALPHGKFDIGFSLPFYKRIINLPVGLK 626
QY 493 DLESVDEBELKSLVWILENDITPV-LDHTFCVHNAPFGRILQHEIKPNRANVPYENKK 551
DB 627 DLESIDPEFINSILWVKNENIEBCEGLMVFSDXKILGEIKSHDLKPNGNILVTEENKE 686
QY 552 EYRLVYNNRFRMGIEAQLALOKGFNELIPQHLKRPDQKLELIIIGLDKIDINDMKS 611
DB 687 EYRMVAMERLSRGVEEQTAFFEGFNEILPQOYLQYFAKLELVLLCGMEIDINDMQR 746
QY 612 NTRLKHCVAADSNIVRMQAVETFEDEERARLIQFTMGSTRVPLQGFALQOSTAGAPR 671
DB 747 HAIYRHYTRTSKQIMFWQFVKEIDNEKMRLLQFTVGTGRLPVGGFADLMGNSN--GPQ 803
QY 672 LFTIHLIDANTDNLPKAHTCFNRIDIPYSEYEKLYEKLTAVEETCGFAVE 723
DB 804 KFCIEKV-GKENWLPKSHTCFNRIDLPYKSYQKLEKLLFAIEETEGFGE 854

RESULT 7

US-09-070-060-3
Sequence 3, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Husted, Carolyn M.
TITLE OF INVENTION: Human E3 Ubiquitin Protein
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENCA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850-5437
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,060
FILING DATE: 30-Apr-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/073, 839
FILING DATE: 05-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Higgins, Patrick H
REGISTRATION NUMBER: 39,709
REFERENCE/DOCKET NUMBER: PHM.70312
TELECOMMUNICATION INFORMATION:

TELEPHONE: 302.886.4889
TELEFAX: 302.886.8221
TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 852 amino acids
TYPE: amino acid
STANDARDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
S-09-070-060-3

Query Match 33.5%; Score 1301; DB 2; Length 852;
Best Local Similarity 34.7%; Pred. No. 5.4e-120;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

Y 6 KIRLVLCAL-NLAKDFFRLDPFAKIVVDGSGQCHSTDTVNTLDPKWNQHYDLYVK 64
b 9 QLOITVISAKLKENKNWFG-PSPYEVTVD--GQSKTEKCNNTNSPKWKQPLTVITP 65
Y 65 TDSITISVNNHKKHK--QGAGFLGCVRLLSNAISRLKDT-----GQOR-----L 108
b 66 VSKLHFRVWSHQTLSKSDVLLGTAALDIYETLKNMKNLEVVVTLQLGSGKEPTETIGL 125
Y 109 DLCKLNPSDPAVRGQIVVSLQT-----RDRTGTGS-----VVDG 144
b 126 SIC---LDGLQLESEVNTGETTCSASQNDGSRKDETRVSTNGSDDEPDAGAGN 181
Y 145 RGL-----LENETVYEDSGPRLSCMEEPAPYDSTGAAGGNCRFVES-----192
b 182 RRVGNNPSLSNGG--FKPSRPPRP-----SRPPPTPRPASVNGSPSATESDGSST 234
Y 193 -----PSQDORLOAQLRNP 208
b 235 GSLPPTNTNTSECATSGLIIPLTISGSGPRPLNPVTOAPLPQWEQKV-----D 286
Y 209 VRGSL-----QTPQNRPHGHSPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP 260
b 287 QGRVYVYDVHVEKRTTWDRPE---PLPPGWERVVDNMGRIVYVDHFTTITWQRTPLE 341
Y 261 -----RDL-----NSVNCDELGLPLPGWEVRSTVSG 286
b 342 SVRYEOMQLORSQLOGAQFNQRFYVGNQDLFATSQSKEFPLGLPLPGWEKRTDSNG 401
Y 287 RIYFVDHNNRTTQFTDRLHMHMHQCOLKEPSQPLP-----324
b 402 RVYFVNHNTRITQWEDPR-----SQQLNE--KPLPEGWEMRFTVDGIPYFVDHNRRT 453
Y 325 -----PSEGLEDDELPAQRYERDLVQKLKVL---HELSQQOQAGHCRIEVSREIFEE 377
b 454 TYIDPRATGSALDNGFOIAYVRDFKAKYQYFRWCQQLAMPQ---HIKITVTRKTLFED 509
Y 378 SYRQIMKMPKDLKRLMVKFRGEGLDYGGVAREWLYLLCHEMLNPPYGLFOYSTDNIY 437
b 510 SFQOIMSFQDLRRRLWVIFPGEGLDYGAVAREWFLLSHEVLNPMYCLPEYAGKDN 569
Y 438 MLQINPDSSINPDHLSYFHFVGRINGLAVFHGYINGGFTVPFYKOLLGKPIOLSLESV 497
b 570 CLQINPASYINPDHLYKFRFGFRTIAMLALFHGKFDITGFSLPFYKILNKPVLKDLSEI 629
Y 498 DPRLHSLWILENDITPV-LDHTCVENHAFGRILQHELKNGRNPVTEENKEIYVL 556
b 630 DPFFYNSLWVKNENIECDLEMYFSVDKEITLKGHDLKPNGNILVTEENKEIYEM 689
Y 690 VAEWRLSRGVEEQTAFFGEPNEILPQQLYQVDAKELEVLICGQWBIIDNDWQRAIYR 749
Y 617 HCVADSNVRFNQAVETDEPRRALLQFTVGTSTRLPQGFALQOSTGAAQPRFLTH 676
b 750 HYARTSQIMWFQVYKEDIDNEKRWLLIQFTVGTGRUPVGVGFADLMGNS---GPKFCIE 806
Y 677 LIDANTDNLFKAHTCFNRLLDIPPEYSEKLYEKLITAVEETCGFAVE 723

Db 807 KV-GRENWLPRSHTCFNRLLDIPPYKSYEQLKXKLLFAIBETEGFQOE 852

RESULT 8

US-09-357-746-3
Sequence 3, Application US/09357746
Patent No. 6087122
GENERAL INFORMATION:
APPLICANT: ZENECA Limited
TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
FILE REFERENCE: PHM.70312.N1
CURRENT APPLICATION NUMBER: US/09/357,746
CURRENT FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: US No. 608712209/070,060
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 852
TYPE: PRT
ORGANISM: Homo sapiens
US-09-357-746-3

Query Match 33.5%; Score 1301; DB 3; Length 852;
Best Local Similarity 34.7%; Pred. No. 5.4e-120;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

QY 6 KIRLVLCAL-NLAKDFFRLDPFAKIVVDGSGQCHSTDTVNTLDPKWNQHYDLYVK 64
Db 9 QLOITVISAKLKENKNWFG-PSPYEVTVD--GQSKTEKCNNTNSPKWKQPLTVITP 65
QY 65 TDSITISVNNHKKHK--QGAGFLGCVRLLSNAISRLKDT-----GQOR-----L 108
Db 66 VSKLHFRVWSHQTLSKSDVLLGTAALDIYETLKNMKNLEVVVTLQLGSGKEPTETIGL 125
QY 109 DLCKLNPSDPAVRGQIVVSLQT-----RDRTGTGS-----VVDG 144
Db 126 SIC---LDGLQLESEVNTGETTCSASQNDGSRKDETRVSTNGSDDEPDAGAGN 181
QY 145 RGL-----LENETVYEDSGPRLSCMEEPAPYDSTGAAGGNCRFVES-----192
Db 182 RRVGNNPSLSNGG--FKPSRPPRP-----SRPPPTPRPASVNGSPSATESDGSST 234
QY 193 -----PSQDORLOAQLRNP 208
Db 235 GSLPPTNTNTSECATSGLIIPLTISGSGPRPLNPVTOAPLPQWEQKV-----D 286
QY 209 VRGSL-----QTPQNRPHGHSPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP 260
Db 287 QGRVYVYDVHVEKRTTWDRPE---PLPPGWERVVDNMGRIVYVDHFTTITWQRTPLE 341
QY 261 -----RDL-----NSVNCDELGLPLPGWEVRSTVSG 286
Db 342 SVRYEOMQLORSQLOGAQFNQRFYVGNQDLFATSQSKEFPLGLPLPGWEKRTDSNG 401
QY 287 RIYFVDHNNRTTQFTDRLHMHMHQCOLKEPSQPLP-----324
Db 402 RVYFVNHNTRITQWEDPR-----SQQLNE--KPLPEGWEMRFTVDGIPYFVDHNRRT 453
QY 325 -----PSEGLEDDELPAQRYERDLVQKLKVL---HELSQQOQAGHCRIEVSREIFEE 377
Db 454 TYIDPRATGSALDNGFOIAYVRDFKAKYQYFRWCQQLAMPQ---HIKITVTRKTLFED 509
QY 378 SYRQIMKMPKDLKRLMVKFRGEGLDYGGVAREWLYLLCHEMLNPPYGLFOYSTDNIY 437
Db 510 SFQOIMSFQDLRRRLWVIFPGEGLDYGAVAREWFLLSHEVLNPMYCLPEYAGKDN 569
QY 438 MLQINPDSSINPDHLSYFHFVGRINGLAVFHGYINGGFTVPFYKOLLGKPIOLSLESV 497
Db 570 CLQINPASYINPDHLYKFRFGFRTIAMLALFHGKFDITGFSLPFYKILNKPVLKDLSEI 629

QY 498 DPELHKSILVILENDIPV-LDHTFCVHNAPGRILOHELKPNRNPVPTVENKEYVRL 556
 DB 630 DPEFYNLSIVKNNIECDLEWVSDKEILGSHDLKNGGNIIVTEENKEEVRM 689
 QY 557 YVWRFWRGTAEQALOKFENLIQHLLKPDQKELELIGGLDKIDNDKSNTRLK 616
 DB 690 VAEWRLSRGVEEQAFEGFENILQQYLOVFDKAELEVLGCGQEIIDLNDQWRAIYR 749
 QY 617 HCVADSNIVFWQAVETFEDEERARLLQVTSRVPLOQGFALQSGTGAGPRLFTIH 676
 DB 750 HYARTSQIWMFQFVEKIDNEXEMLLQFVIGTCKLPVGGFADLMGNN---GPKFCIB 805
 QY 677 LIDANTNLKANTCFNRIDIPYVESYKLYKLLTAVETTCGRAVE 723
 DB 807 KV-GKENWLFASHTCFNRLLDPLPYKSYEQKELLLFAIEETEGFQGE 852

RESULT 9

US-08-895-601-6
 : Sequence 6, Application US/08895601
 : Patent No. 6060262

GENERAL INFORMATION:

APPLICANT: Beer-Romero, Peggy
 APPLICANT: Strack, Peter J.
 APPLICANT: Glass, Susan J.
 APPLICANT: Rolfe, Mark

TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,
 TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
 NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
 APPLICATION NUMBER: US/08/895,601

FILING DATE: 16-JUL-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: MIV-096.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 927 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

-08-895-601-6

Query Match

Best Local Similarity

Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

32.4%; Score 1259; DB 3; Length 927;

33.6%; Pred. No. 9.4e-116;

7 IRLTVLCAKHLAKDFFRLPDPFAKIV-----VDGSGQCHSTDVNTVNLDPKWN----- 55

48 VVRVTVAGIAGLAKKIDLGASDPYRVTLVDPNMGVLTSTVQTKIKSLNPKWNEILFRV 107

56 ---QH-----YDLVYGVKTD--SITISVWN-----HKKHKQOG 83

108 HPQCHRLFEVFDENLRTRDDFLGQVDVPLXPLPTENPRLERPPTFXDFVLHPRSHKSRV 167

QY 84 AGFLGCVRLLSNAISLKD-----GYQRLD-----LCKL----- 113
 DB 168 KGYL-----RLKMTYLPKTSGEDDVAEQAELFPGWVLDOPDAAACHLQOOQE 216
 QY 114-----NPSD-----TDAVRQIVVLSQ-----TRDR 134
 DB 217 PSLPPQWEERQDILGRYVYVNHESRRRTQWKPTQDNLTDANGNIQLQACRAFTTRQ 276
 QY 135 IGTGGSVVDCRGLLEN-----EGTVYEDSG-PGRPLSCFMEEPAPYTD----- 176
 DB 277 ISETESVDNQESSENWEIREDATMYSSQAFPPPPSSNLDVPTHLAELNARLPIFG 336
 QY 177-----STGAAAGGNCR-----FVESPS-----QDORLQAQRL-- 204
 DB 337 NSAVSQPASSNSHRSRGLQAYTFEEQPTLPVLLPTSSGLPPGWEEKQDGRGSIYYVDH 396
 QY 205-----RNPDRVRSGLQTPQ-----NRPHGHQSPPELPEGEYEQ 234
 DB 397 NSRTTWTWKPTVQATVETSQTSQSSAGSQASTSDSGQVQTPSPSEIQGLPKGWEV 456
 QY 235 RTTVQGVYELHTQTGVSTWHDPR--IPDLNSV-----NCDELGPLPGWEVSTVSGRI 288
 DB 457 RHPNGRPFIDHNTKTTTWEDPRLKIPAHLEKGTSLDTSNDLGLPLPGWEERTHTDGR 516
 QY 289 YFVDHNNRTTQFTDPRLHHMHQCOLKEPSQFLPLPSEGSELEDEELPAQRYERDLVQKL 348
 DB 517 FYINHNKRTOWEDPLENV-----AITGPAVP-----YSRDYKRY 553
 QY 349 KVLHLSLQOPQAGHCRIVRSREBIFPESYRQIKMTPKD-LKKRLMVKTRGEGLDYG 407
 DB 554 EFRKRLKQNDIPNKFEMKLRATVLEDSYRRIMGVRRADFLKARLWIEPDGSGLDYG 613
 QY 408 GVAREMLYLLCHEMLNPYYGLFYQYSTDNIYMLQINPDSS1-NPDHLSYFHFVGRIMGLAV 466
 DB 614 GVAREMFLISKEMFNYYGLFEYSATONYTLQINPNSGLCNEHLSYKFIQVAGNAV 673
 QY 467 FHGHYNGGFTVPFYKOLLGKPIQLSDLESVDPELHSLVILENDITPVLDTHTFCVHN 526
 DB 674 YHGKLDGFFIRFYQMLHKFITLHDMESVDSEYNSLRWILENDPTE-LDLRFLIDEE 732
 QY 527 AFRILQHELKPNRNPVPTVENKEYYRLVYVNRFMGIEAQAFLAQKGNELIPQHLL 586
 DB 733 LFGQTHQHELKNGSBIYVTVNKNKXEYIYLVQWRFVNRIOQMAAFKEGFFELIPQDLI 792
 QY 587 KPFDQKLELLIGGLDKIDLNDKSNTRLKH-CVADSNIVFWQAVETFEDEERARLIQ 645
 DB 793 KIPDENELELMCGLGDVDVNDREHTKYKGYSHQVIOFWFRAVLMDMDSEKRIQLQ 852
 QY 646 FVTGSTRVPLQGFALQSGTGAGPRLFTIHLIDANTNLKPAHTCFNRIDIPPYSEYK 705
 DB 853 FVTGSTRVPMNGFAELYGSN---GQSFTEQW-GTPEKLPRAHTCFNRDLDPYSEFEE 908
 QY 706 LYEKLLTAVEETCGF 720
 DB 909 LWDKQLQWAIENQGF 923

RESULT 10

US-08-539-205A-6
 : Sequence 6, Application US/08539205A
 : Patent No. 6001619

GENERAL INFORMATION:

APPLICANT: Beach, David H.
 APPLICANT: Caligiuri, Maureen

APPLICANT: Nettek, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-539-205A-6

Query Match 31.6%; Score 1228.5; DB 3; Length 834;
Best Local Similarity 40.4%; Pred. No. 8.5e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

93 LSNALSRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
186 LSEELSRRLQITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQHLPFSPVAVYHTTTPG 245
135 IGTGGSV-VDCRGLL-----ENEGTVVDSGCRPLSCFWEPPAPYTDSTGAAGGN-- 186
246 LPSGWEERKDAKGRYYVNNRRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 295
187 -----CRFVESPSQORLOAQRLNPDVRGSL-----QTPQNRPHGOSPE----- 227
296 IEPQIRPRSLSSPTVTLAPLEGAKDSPVRAVKDTLGNPQSPSPYNSPKQHKVTQ 355
228 --LPEGYQRITVQGVYFLHTQTGVSTMDPRIP-----ROLNSVNCDELGLPLPGWEV 280
356 SFLPGWEMRIAPNGRPFIDHNTKTTTWDPRLKPFVHMRKSTLNPNDLGLPLPGWEE 415
281 RTVSGRIYFVDHNNRTQFTDPRLHHIMHQCLKEPSQPLPSEGSLEDEELPAQRY 340
416 RIHLDGRFTYDHNKSKITQWEDPRLOQ-----PAITG-----PAVY 452
698 PYESYKLYEKLTLTAVETCGF 720

31.6%; Score 1228.5; DB 4; Length 834;
Best Local Similarity 40.4%; Pred. No. 8.5e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

93 LSNALSRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
186 LSEELSRRLQITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQHLPFSPVAVYHTTTPG 245
135 IGTGGSV-VDCRGLL-----ENEGTVVDSGCRPLSCFWEPPAPYTDSTGAAGGN-- 186
246 LPSGWEERKDAKGRYYVNNRRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 295
187 -----CRFVESPSQORLOAQRLNPDVRGSL-----QTPQNRPHGOSPE----- 227
296 IEPQIRPRSLSSPTVTLAPLEGAKDSPVRAVKDTLGNPQSPSPYNSPKQHKVTQ 355
228 --LPEGYQRITVQGVYFLHTQTGVSTMDPRIP-----ROLNSVNCDELGLPLPGWEV 280
356 SFLPGWEMRIAPNGRPFIDHNTKTTTWDPRLKPFVHMRKSTLNPNDLGLPLPGWEE 415
281 RTVSGRIYFVDHNNRTQFTDPRLHHIMHQCLKEPSQPLPSEGSLEDEELPAQRY 340
416 RIHLDGRFTYDHNKSKITQWEDPRLOQ-----PAITG-----PAVY 452
698 PYESYKLYEKLTLTAVETCGF 720

31.6%; Score 1228.5; DB 4; Length 834;
Best Local Similarity 40.4%; Pred. No. 8.5e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

93 LSNALSRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
186 LSEELSRRLQITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQHLPFSPVAVYHTTTPG 245
135 IGTGGSV-VDCRGLL-----ENEGTVVDSGCRPLSCFWEPPAPYTDSTGAAGGN-- 186
246 LPSGWEERKDAKGRYYVNNRRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 295
187 -----CRFVESPSQORLOAQRLNPDVRGSL-----QTPQNRPHGOSPE----- 227
296 IEPQIRPRSLSSPTVTLAPLEGAKDSPVRAVKDTLGNPQSPSPYNSPKQHKVTQ 355
228 --LPEGYQRITVQGVYFLHTQTGVSTMDPRIP-----ROLNSVNCDELGLPLPGWEV 280
356 SFLPGWEMRIAPNGRPFIDHNTKTTTWDPRLKPFVHMRKSTLNPNDLGLPLPGWEE 415
281 RTVSGRIYFVDHNNRTQFTDPRLHHIMHQCLKEPSQPLPSEGSLEDEELPAQRY 340
416 RIHLDGRFTYDHNKSKITQWEDPRLOQ-----PAITG-----PAVY 452
698 PYESYKLYEKLTLTAVETCGF 720

b 453 SREFKQYDYFEKLLKPADIPNRFEMKLNHNHIFESSYRINSKVPDVLKARLWIEPE 512
y 400 GREGLDYGGVAREWYLICHEMLNPIYGLFOYSTDNIMYLMQINDPSI - NPDHLSYFHFV 458
b 513 SKGGLDYGGVAREWFLSKEMFNFPYGLFYSATDNYTLQINPNSGLCNEDHLSYFTFI 572
y 459 GRIMGLAVPHGHYINGGFTVPFYKOLLGKPIQLGSLSDVDELHKSLLWILENDITPVL 518
b 573 GRVAGLAVFHGLLDGFFIRFPYKMLGKQITLNDMESVSEYNSLKNWILENDITE-LD 631
y 519 HTPCEVHNAGRILOHEUKPKGNRPVTEENKGYRLYVNRFRMGIEAQFIALOKGFN 578
b 632 LMFCDEENFGQTYQDLKPNKSEIMVTNENKREYIDLVIQWRFVNRVQKMAFLEGFT 691
y 579 ELIPOHLLKPPDKLELILIGLDKIDLNDKSNTRLKHCVADSN - IVRNFWOAVETFE 637
b 692 ELLPDLIKIFDENLELLMCGLDVNDNRQHSIYKNGYCPNHPVIQWFKAVLLMDA 751
y 638 ERRARLLQVTSRVPVQGGFKALGGSTGAAGPRLFTHLIDANTMLPKAHTCFNRIDI 697
b 752 EKRIQLQVTSRVPVPMNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRIDL 807
y 698 PYVESYKLYEKLTAVEETQCF 720
b 808 PPTETFEDELREKLLMAVENAQGF 830

RESULT 12
S-08-630-916A-48
Sequence 48, Application US/08630916A
Patent No. 601137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
NUMBER OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELEPHONE: (212) 750-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
IS-08-630-916A-48

Query Match 31.5%; Score 1225; DB 3; Length 906;
Best Local Similarity 35.5%; Pred. No. 2,2e-112;
Matches 281; Conservative 115; Mismatches 227; Indels 169; Gaps 19;

QY 28 PFAKIVDGSQCHSTDTVKNLTDPKWNQHY-----DLVYVKTDSTITISVWNHKKIHKQG 83
DB 188 PNGSALTGDS-QLPSESDSSGTAPE-NRHQPSTNCFGGRS-----RTHRHSG 234
QY 84 AGFLGCVRLLSNAISRKLDGTGYQRLDLCKLNPSDSDTDAVRGQIVSVLQTRDR--IGTGGSV 141
DB 235 AS-----ARTTPATGEQSPGARSRRHROPVKNSGHS 264
QY 142 VDCGLLENGTVEED-----SGQRPUSCFMEE-----PAPYTDSTGAAAGGN 186
DB 265 GLANGTVNDEPTTATDPEEPSVVGVTSPPAAPLSVFPNFTTSLPAPATPAEG-----317
QY 187 CRFPESQD--ORLQAOURLNPDVRGSLQTPQNRPHGHSPELPE-----230
DB 318 ---EESTSTQQLPA-----AAQAPALPAGWEQRELPNGRVVYVDHNTKTTT 363
QY 231 ---GYORTTQQQVYFLHTQTGCVSTWHDPRIPRDLN-----264
DB 364 WERPLPQWEKRTDPRGRFYVDHNTTTRTTTQRTAAYVRYVNEQMSQORNOLQGANQHFS 423
QY 265 ---SVNCDLGLPLPGWEVRSTVSGRIYFVDHNNRTTQFTDPR-----304
DB 424 QRFLYQFWSASTDHDPLGPPGWEKRD--NGRVYVNTNTRTTQWEDERTQGMIOEPAL 482
QY 305 ---LHHMNHQCLKEPSQPLPSESGSLEDELPQAQRYERDLVQKLKVL 352
DB 483 PPGWEMKYTSEGVRYFVDHNTTTRTTKDPKPPGPGESGKQSGPGAYDRSPRWKYHQFRPLC 542
QY 353 HELSLQPOAGHCHIEVSREIPEESYRQIMKMRPKDLKXLMVKFRGEGLDYGGVASE 412
DB 543 HSNAL----PSHVKISVSROTLPFEDSFQIMNMKPYDLRRRLYIIMRGEGLDYGGIAIE 598
QY 413 WLYLLCHEMLNPPYGLFQYSTDNIMYLMQINPDSSINPDHLSYFHFVGRIMGLAVFHGYI 472
DB 599 WFFLLSHEVLNPMYCLEYAGKNNYCIQINPASSINPDHLYFRFGRPTAMALYHGKFI 658
QY 473 NGGTVVPYKOLLGKPIQLSDLSVDDELHKSLLWILENDITPV-LDHTPCVEHNAFGRI 531
DB 659 DTGFTLFPYKMLNKRPTLKDLSIDPEFYNSIVTKENNEECGLELYFIQDMEILGV 718
QY 532 LQHELKPNRNPVPTENKKEYVRLYVNMRFMEGIEAQFLALQKGFNELIPQHLKPFQ 591
DB 719 TTHELKEGESIRVTEENKKEYIMLTDRFTGVEEQTKAFLDGNEVAPLEWLVFDE 778
QY 592 KELELIIGLDKIDLNDKSNTRLKHCVADSNIVRWFQAVTTFDEERRARLLQFTVGST 651
DB 779 KELEMLCGMOEIDMSDWKSTIYRHYTKNSKQIQWFWQVVKEMDNKRIQLLQFTVGTG 838
QY 652 RVPLQGGKALQSGTGAAGPRLFTHLIDANTMLPKAHTCFNRIDIPPYESYKLYEKL 711
DB 839 RLPVGGFAELIGSN---GPQKFCIDKVKET-WLPRSHTCFNRDLDFPYKSYZOLREKL 894
QY 712 TAVEETCGFAVE 723
DB 895 YAIETEGFGQE 906

RESULT 13
US-08-630-916A-46
Sequence 46, Application US/08630916A
Patent No. 601137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
NUMBER OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York

RESULT 14
;-08-247-904B-8
Sequence 8, Application US/08247904B
Patent No. 5981699

```

Query Match      13.7%; Score 530.5; DB 2; Length 874;
Best Local Similarity 26.1%; Pred.No. 2.7e-43;
Matches 190; Conservative 117; Mismatches 262; Indels 159; Gaps 26

QY      98 SRLKDTGYRLDLCKLNPSTD---DAVRQIVSVISLQTRDRICTG-----GSVVDCC 145
Db      202 SRIGSSQGDNNLQKLGPDVDSVIDAIR-RVYTRLLSNSKEIETAFNLVALVYLSPNVCCD 260

QY      146 GLENEGTVY-----EDSGEGRPLSCFMEEPAPYTDSTCGAAG 183
Db      261 LTVHN---VYSRDPNYLNLFTIGMENENLHSPLEYLMAFLPCAKMSKLP-----LAAQ 311

QY      184 GGNCRFVESPSDQ-----RLQQRLRNPD----- 208
Db      312 GKILRLSKYNADQIRMMETFQQLITYKVISNEFNLSRNLVNEFNSNLVNDDDAIVAAS 371

QY      209 -----VRGSLQTPQNRPHGHOS-PELPE-----GYEQRTTVGGQVYFLHTQT 249
Db      372 KCLQWYIYANVVGGEVDTHNNEEDDEPIPESSBELTLQELLGEERNKKGLRYDPLETEL 431

QY      250 GYST--WHDPRIPRD--LNSVNCDELGPLPGGNEVRSTVSGRIYFVDHNNRTTQFTDPL 305
Db      432 GWKTLDCRKPLIPPEEFINE-----PLNEVLMDKDYT--FFKVETENKSPFMTCPFI 482

QY      306 HHIMNHQCLKEPSQFLEPLPSEGSLEDEELPAQRYERDLVQKLVIJRHLSLQQPOQAGHC 365
Db      483 LNAVTKNLGLYYDNR-IRWYSE-----RRITVL-YSLVQGOQLNPYL 522

QY      366 RIVSREEEIFEES-----YRQIKMRPKDLKKELMVKFGEESGLDYGVAERMLVLLCHEML 422
Db      523 RLKVRDRDH1IDDALVRLEMIAMENPADLKKQLYVEFEGEGDEGGVSKFEFFGLVVEEIF 582

QY      423 NPYIGLGFQYSTDN1YMLQINPDSSINPDHLSYFHVGRINGLAVFPHGHTYINGGTFVPPYK 482
Db      592 NPIGNMFTWYD-ESTVLYENFENSSP-----ETEGQFTIIGIVLGLAIYNNCLLDVHFPMWYVR 638

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on on: February 20, 2004, 15:27:02 ; Search time 31.9477 Seconds
(without alignments)
4738.482 Million cell updates/sec

tie: US-10-009-945-2
rfect score: 3884
quence: 1 GSSIKIRLTVLCAKWLAKK.....EKLYEKLITAVEITCGFAVE 723

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 801455 seqs, 209382283 residues

tal number of hits satisfying chosen parameters: 801455

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

tabase :

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3878	99.8	722	15	US-10-097-534-14
2	3001	77.3	748	12	US-10-021-660-81
3	2913.5	75.0	735	12	US-10-313-955-2
4	2236	57.6	514	12	US-09-764-875-819
5	1471	37.9	766	12	US-10-313-955-4
6	1436.5	37.0	832	12	US-10-032-585-7296
7	1345	34.6	869	15	US-10-128-714-8162
8	1259	32.4	927	15	US-10-097-534-15
9	1249.5	32.2	911	15	US-10-208-823-279
10	1245.5	32.1	739	12	US-10-374-979-89
11	1245.5	32.1	739	15	US-10-097-534-10
12	1245.5	32.1	752	11	US-09-919-039-235
13	1228.5	31.6	834	12	US-10-313-955-6
14	1228.5	31.6	995	15	US-10-097-534-9
15	1228.5	31.6	995	15	US-10-205-823-275

16	1227	31.6	725	15	US-10-185-050-126	Sequence 126, App
17	1226.5	31.6	759	15	US-10-128-714-3162	Sequence 3162, App
18	1226	31.6	854	15	US-10-205-823-277	Sequence 277, App
19	1225	31.5	870	15	US-10-097-534-12	Sequence 12, Appl
20	1225	31.5	906	15	US-10-185-050-48	Sequence 48, Appl
21	1223	31.5	854	12	US-10-287-218-3	Sequence 3, Appl
22	1212	31.2	474	11	US-09-774-639-371	Sequence 371, App
23	1212	31.2	474	11	US-09-989-730-249	Sequence 249, App
24	1047	27.0	683	15	US-10-185-050-46	Sequence 46, Appl
25	1047	27.0	684	15	US-10-097-534-11	Sequence 11, Appl
26	1023	26.3	380	15	US-10-307-956-32	Sequence 32, Appl
27	1012	26.1	375	15	US-10-097-534-13	Sequence 13, Appl
28	1002	25.8	1562	15	US-10-043-487-300	Sequence 100, App
29	836.5	21.5	1094	15	US-10-043-487-285	Sequence 285, App
30	836.5	21.5	1488	15	US-10-097-534-16	Sequence 16, Appl
31	821.5	21.2	733	15	US-09-925-300-1628	Sequence 1628, Ap
32	724	18.6	277	10	US-10-268-036-5	Sequence 5, Appl
33	720.5	18.6	358	15	US-09-925-300-1527	Sequence 1527, Ap
34	677.5	17.4	276	10	US-09-764-870-384	Sequence 384, App
35	612	15.8	125	9	US-09-764-870-384	Sequence 1110, Ap
36	612	15.8	125	12	US-09-764-875-1110	Sequence 384, App
37	612	15.8	125	15	US-10-125-340-384	Sequence 25, Appl
38	518	13.3	823	15	US-10-097-534-25	Sequence 24, Appl
39	449.5	11.6	1050	15	US-10-097-534-24	Sequence 28, Appl
40	449.5	11.6	1054	15	US-10-097-534-29	Sequence 29, Appl
41	449.5	11.6	1024	15	US-10-097-534-30	Sequence 30, Appl
42	392	10.1	584	12	US-10-108-260A-4173	Sequence 4173, Ap
43	389	10.0	1022	12	US-10-287-218-16	Sequence 16, Appl
44	389	10.0	1022	15	US-10-286-036-2	Sequence 2, Appl
45	389	10.0	1022	15	US-10-286-036-2	

ALIGNMENTS

RESULT 1
US-10-097-534-14
; Sequence 14, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-534-14

Query Match 99.8%; Score 3878; DB 15; Length 722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 722; Conservative 0; Mismatches 0; Gaps 0;

QY 2 GSSIKIRLTVLCAKWLAKKDFRLPDPFAKIVDVGSGCHSTDTVKTLDPKMNHVDLY 61
Db 1 GSSIKIRLTVLCAKWLAKKDFRLPDPFAKIVDVGSGCHSTDTVKTLDPKMNHVDLY 60
QY 62 VGKTDSTITISVANKHKIKKOGAGPLGCVRLISNAISRLKDTGYQRDLCKLNPSDPAV 121

61 VGTDSITISVWVHKKHKKQAGAGLGCYVRLLSNALSRKDTGYORLDLCKLNSPDTDAV 120
122 RGOIVSVLQTRDRIGTGGVWDCRGLLENEGTVDSPGRPLSCFMEEPAPYTDSTGAA 181
121 RGOIVSVLQTRDRIGTGGVWDCRGLLENEGTVDSPGRPLSCFMEEPAPYTDSTGAA 180
182 AGGNCRFVESPQDORLOARLNPVDRGSLQTPQNRPHGHQSPDPGEGYEQRTTVQGO 241
181 AGGNCRFVESPQDORLOARLNPVDRGSLQTPQNRPHGHQSPDPGEGYEQRTTVQGO 240
242 VYFLHTQTVSTWHDPRIPRDLSNVCNDBELGPPGWEVSTVSGRIYFVDHNNRTTQFT 301
241 VYFLHTQTVSTWHDPRIPRDLSNVCNDBELGPPGWEVSTVSGRIYFVDHNNRTTQFT 300
302 DPLHLMHMQCOLKEPSPQPLPSEGSLDEELPAQYERDLVQKLVLRHLSLQOPQ 361
301 DPLHLMHMQCOLKEPSPQPLPSEGSLDEELPAQYERDLVQKLVLRHLSLQOPQ 360
362 AGHCRIEVSREEIFEBSYRQIMKMRPKDKKRLMVKFRGEGLDYGGVAREWLYLLCHEM 421
361 AGHCRIEVSREEIFEBSYRQIMKMRPKDKKRLMVKFRGEGLDYGGVAREWLYLLCHEM 420
422 LNPYYGLFOYSTDNIMYLQINPDSSINPDHLSYFHFVGRINGLAVFHGHYINGGFTVPFY 481
421 LNPYYGLFOYSTDNIMYLQINPDSSINPDHLSYFHFVGRINGLAVFHGHYINGGFTVPFY 480
482 KOLLGKPIQLSDLESVDPELHLSLVWILENDITPVLDHTFCVHNFAFRILQHELKFNGR 541
481 KOLLGKPIQLSDLESVDPELHLSLVWILENDITPVLDHTFCVHNFAFRILQHELKFNGR 540
542 NVPTEENKEVYVRLVNVNRFMRGTEAQFLALOKGFNBLIQHLLKPEDQXELIIGGL 601
541 NVPTEENKEVYVRLVNVNRFMRGTEAQFLALOKGFNBLIQHLLKPEDQXELIIGGL 600
602 DKIDLNDKSNTRLKHCVDADSNIVRFWQAVTFDEERRARILQFVTGSTRVPLQGFKAL 661
601 DKIDLNDKSNTRLKHCVDADSNIVRFWQAVTFDEERRARILQFVTGSTRVPLQGFKAL 660
662 QSGTGAAGRLFTIHLIDANTDNLKPAHCTFNRIIDIPYSEYKLYEKLTAVEETCGFA 721
661 QSGTGAAGRLFTIHLIDANTDNLKPAHCTFNRIIDIPYSEYKLYEKLTAVEETCGFA 720

SULT 2

-10-021-660-81
Sequence 81, Application US/10021660
Publication No. US20030152926A1
GENERAL INFORMATION:
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.
APPLICANT: EOS Biotechnology, Inc.
TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
FILE REFERENCE: 018501-000710US
CURRENT APPLICATION NUMBER: US/10/021,660
CURRENT FILING DATE: 2001-12-06
PRIORITY FILING DATE: 2001-12-06
PRIORITY FILING DATE: 2001-02-14
PRIORITY FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 135
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 81
LENGTH: 748
TYPE: PRT
ORGANISM: Homo sapiens
-10-021-660-81

Query Match 77.3%; Score 3001; DB 12; Length 748;
Best Local Similarity 74.6%; Pred. No. 6e-266;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;
QY 5 IKIRLTVLCAKNLAKKOPFRLPDPFAKIVVDDGSGQCHSTDTVKNLTDPKWNQHYDLYVGK 64
DB 12 VKLRLTVLCAKNLAKKOPFRLPDPFAKIVVDDGSGQCHSTDTVKNLTDPKWNQHYDLYIGK 71
QY 55 TDSITISVWVHKKHKKQAGAGLGCYVRLLSNALSRKDTGYORLDLCKLNSPDTDAVQGO 124
DB 72 SDSVITISVWVHKKHKKQAGAGLGCYVRLLSNALSRKDTGYORLDLCKLNSPDTDAVQGO 131
QY 125 IVVSLQTRDRIGTGGVWDCRGLLENEGTVDSPGRPLSCFMEEPAPYTDSTGAA 155
DB 132 IVVSLQTRDRIGTGGVWDCRGLLENEGTVDSPGRPLSCFMEEPAPYTDSTGAA 191
QY 156 --EDSGPRLSCFMEEPAPYTDSTGAAAGGNCRFVESPQDORLOARLNPVDRGSL 213
DB 192 ASEYSSPGRPLSCFMEEPAPYTDSTGAAAGGNCRFVESPQDORLOARLNPVDRGSL 242
QY 214 QTPQNRPHGHQSPDPGEGYEQRTTVQGOVYFLHTQTVSTWHDPRIPRDLSNVCNDBELG 273
DB 243 ---SRTHLHTPPDLPPEGYEQRTTVQGOVYFLHTQTVSTWHDPRIPRDLSNVCNDBELG 298
QY 274 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPR---LHIMNHQCOLKEPSPQPLPSEGS 329
DB 299 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPR---LHIMNHQCOLKEPSPQPLPSEGS 354
QY 330 L---EDELPQAQYERDLVQKLVLRHLSLQOPQOAGHCRIEVSREEIFEBSYRQIMKMR 386
DB 355 LCPDTECLTVPRKRDVLVQKLVLRHLSLQOPQOAGHCRIEVSREEIFEBSYRQIMKMR 414
QY 387 PKDLKRLMVKFRGEGLDYGGVAREWLYLLCHEMNPYYGLFOYSTDNIMYLQINPDSS 446
DB 415 PKDLKRLMVKFRGEGLDYGGVAREWLYLLCHEMNPYYGLFOYSTDNIMYLQINPDSS 474
QY 447 INPDHLSYFHFVGRINGLAVFHGHYINGGFTVPFYKOLLGKPIQLSDLESVDPELHLSLV 506
DB 475 VNPTEENKEVYVRLVNVNRFMRGTEAQFLALOKGFNBLIQHLLKPEDQXELIIGGL 534
QY 507 WILENDITPVLDHTFCVHNFAFRILQHELKFNGRNVPTEENKEVYVRLVNVNRFMRGI 566
DB 535 WILENDITPVLDHTFCVHNFAFRILQHELKFNGRNVPTEENKEVYVRLVNVNRFMRGI 594
QY 567 EAQFLALOKGFNEVLPQHLKPPDQKELELIIGGLKIDNDKSNTRLKHCVDADSNIVR 626
DB 595 EAQFLALOKGFNEVLPQHLKPPDQKELELIIGGLKIDNDKSNTRLKHCVDADSNIVR 654
QY 627 WFOQAVETFEERRARILQFVTGSTRVPLQGFKALQSGTGAAGRLFTIHLIDANTDNL 686
DB 655 WFOQAVETFEERRARILQFVTGSTRVPLQGFKALQSGTGAAGRLFTIHLIDANTDNL 711
QY 687 KAHTCFNRIDIPYSEYKLYEKLTAVEETCGFAVE 723
DB 712 KAHTCFNRIDIPYSEYKLYEKLTAVEETCGFAVE 748

RESULT 3

US-10-313-955-2
Sequence 2, Application US/10313955
Publication No. US20030199036A1
GENERAL INFORMATION:
APPLICANT: Beach, David H.
Nefsky, Bradley
Caligiuri, Maureen
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA

SULT 5

-10-313-955-4

Sequence 4, Application US/10313955

Publication No. US20030199036A1

GENERAL INFORMATION:

APPLICANT: Beach, David H.

Caigiuri, Maureen

Neskey, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/09/392,163

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

PUBLICATION DATE: 05-Dec-2002

150 NEGTV-YEDSGPGR---PLSCFMEEPAPVTDSTGAAGGNCRFVESPSQDORLQARLR 205
192 SDATVANGSGPTSSLPPTCQHPETAATPGGAAGAASQYSSPFDQYGRLLPPGWER 251
206 NPDVRG-----SLQTPQNRPHGHOSP----- 226
252 RTDNFGSTYYVDHNSRNTTWRPALHQSETERGQORQSQSEAEARRQHRGRTLPGEGSVSP 311
227 -----ELPEGYEORTVQGV 242
312 LPTGSGNSITSGNVTVNASANTPVNPAASVMAASGATSGELGELPSGWEQRTFTEGRP 371
243 YFLHTQFVSTWHDPRIPRLDLSVN-----CDELGPLPPGWEVSTVSGRIYFVDH 293
372 YFVDHNRTRTTTWDPRRQYVIRTEGENTTQQQVPSQLGFLPSGWEMLTNTARVYFVDH 431
294 NNRTQFTDRLHHIMNHQCLKEPQPLPSEGSLEDEELPAQRYERDLVOKLVLRH 353
432 NTKTTWDDPR-----LPS--SL-DQNVF--QYKRDPRKRYIFRS 467
354 ELSQQPQAGHCHRIEVSREBIFESYRQIMKRPKDLKRLMKVFRGEGLDYGGVAREW 413
468 QPAL-RILPQOCHIKVRDHIPEDSYQEIQRQTPEDLKKELMIKFDGEGLDYGGVAREF 526
414 LYLICHEMLNPNYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAVFHGIN 473
527 FFLSHDMFNFYCLFEYSHDNYTTLQINPNSGINPEHLNYPKFIGRVVGLGVFRRFLD 586
474 GGFTVPYKOLLKPIQLSDLESVDPELHKSILVILENDITPVLDTFCVENAFGRILQ 533
587 AFFGALYKWLHKVQLQDMEGVDAEFYRSLKWLNDITGILDITSABESFEIVE 646
534 HELKPNRNVPTVEENKEVRYLVNWRMRGLEAQLALQGFNELIPQHLKPFDOKE 593
647 VDLKPGGRDIEVTEENKHEVLEITERISKREVEQKAFIDGFNELIPQELVNVFDSRE 706
594 LELIIGLDKIDLNDKSNRLKHCADVNSIVFWQAVETDEERRARLLOFVGTSTRV 653
707 LELLIGLSAIEDCKWKHDYGYQENDCVIOWFWKCNEDWSQKARLLOFTGTSTRI 766
654 PLOQFKALQSGTAGPRLFTIHLIDANTDNLPAKTCNRIDIPYSEYKLYEKLTA 713
767 PVNGFKDLOGSD--GPRRTTIEKA-GEANQLPKSHTCFNRVLDLPPYTDYESLKQKTLA 822
714 VEETCGFAVE 723
823 VEETVGFQGE 832

RESULT 7
Sequence 8162, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wenqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroskin, Alexey M
APPLICANT: Lenieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIORITY FILING DATE: 2001-04-23
PRIORITY FILING DATE: 2001-04-23
PRIORITY FILING DATE: 2001-04-23
PRIORITY FILING DATE: 2001-04-27
PRIORITY FILING DATE: 2001-06-05
PRIORITY FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: Patent in version 3.1
SEQ ID NO 8162
LENGTH: 869
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-8162
Query Match 34.6%; Score 1345; DB 15; Length 869;
Best Local Similarity 35.9%; Pred. No. 7e-114;
Matches 317; Conservative 123; Mismatches 236; Indels 208; Gaps 26;
QY 11 VLCAKILAKKDFR-----LPDPFAKIVVDGSGOCHSDDTKVNT 49
DB 23 VTAADGLYKRDVFRKSYLAILLVVRILTSGAFAGPDPFAVAVGGE-QTHHTSVIKKT 81
QY 50 LDPKMNQHYDLVVGKTDSTITISVNNHKKIHKQAGFLGCVRL-- 93
DB 82 LNPYNMEMFDMRVNEDSILAIQIFDQKXF-KKQDQGLGVINVRIGDIVDLQMGDGE 140
QY 94 ---SNAISRLK---DTGVQRDLCLKLPSTDAVGOIVVSLQ-- 131
DB 141 PIRHSCDVSRLLFRLLLETEMLTRDLKSN--DNLVVGKLIINLSTNLSTPNTNOANG 198
QY 132 --RDRIGTGGVDCRGLLEN-----EGTYED---SGFG-- 166
DB 199 LHRSHVQSSTS---SGLVPQVAPSSHPAASGAPVDPASNPSLNPQRPVSTPSST 254
QY 167 FWEIPA-----PYDSTGAAGG-----GNCRFE-----SP 193
DB 255 AAPASAGAASVNSHSGSRNLSSPDSQGRLPAGWERREDNLGRTYVVDHNRTRTTWTRP 314
QY 194 SDQRLQAQRLN-----PDVR---GSLQTPQNRPHGHOSP----- 226
DB 315 SSNYNEHAQRSREANMQLEREAHOSRMLPEDRTGANSPLPSSQQAHTFPAGGSANAV 374
QY 227 -----ELPEGYEQRTTVQGVYFLHTQTVSTWHDPR-----IPRDLNSV 266
DB 375 SMMATGATTAGTGELPPGWEQRTTPEGYPYFVDHNRTRTTTWDPRRQYIRMYQONANGT 434
QY 267 N-----CDELGPLPPGWEVSTVSGRIYFVDHNRRTQFTDRLHHIMNHQCLKEPS 319
DB 435 NTTIQQQPVSQLGPLPSGWEMLTNTARVYFVDHNTKTITWDDER----- 479
QY 320 QPLPSPSGSLEDEELPAQRYERDLVOKLVRLHLSLQOPQAGHCHRIEVSREBIFESY 379
DB 480 ---LPS--SL-DQGV--QYKRDPRKLYIFRSQPAL-RIMSGQCHYKVRNNIFEDSY 529
QY 380 ROIMKMRPKDLKRLMKVFRGEGLDYGGVAREWLYLLICHEMLNPNYGLFOYSTDNIMYL 439
DB 530 AEIMRQASDLKRLMIKFDGEDGLDYGSGEPFFLLSHENFNPFFCLFEYSADNYTL 589
QY 440 QINPDSSINPDHLSYFHFVGRIMGLAVFHGYINGGFTVPYKOLLKPIQLSDLESVDP 499
DB 590 QINPHSGVNPHELNYPKFTIGRVVGLAI FHRRLDLSFFTFGAFYKMLKRYSLQDMEGVDE 649
QY 500 ELHKSIAVLENDITPVLDTHTPCVEHNAFGRILQHELKPNCRNVPTVEENKEVRYLVN 559
DB 650 DLHRNUTWTDNDIEGVLELTPSDDERKGERRTIDLPAGGRDIPVNNENKAEVRLVTE 709
QY 560 WRFMRGIEAQFLALQGFNELIPQHLKPFDOKELELIGGLDKIDLNDKSNTRLKHCV 619
DB 710 WKIVKRVEEQFNAPFSGFNELIPADLVNVFDERELELLIGGIADIDVDDKKHTDYGQ 769
QY 620 ADSNIVRWQAVETFEERRARILQFTVGTSTRVPLQGFKALQSGTAGPRLFTIHLID 679
DB 770 ESDEVIONFWKIVRSWDAEQSKRLLOFTGTSTRIPVNGFKDLOGSD--GPRRTTIEK-S 825
QY 680 ANTDNLPAKHTCFNRIDIPYSEYKLYEKLITAVEETCGFAVE 723
DB 826 GDPALPKSHTGTGNRLDLPYKSYETLEHKMSIAVEETLGFQGE 869

```
RESULT 8
3-10-097-534-15
Sequence 15, Application US/10097534
Publication No. US20030049607A1
GENERAL INFORMATION:
APPLICANT: GREENER, TSVIKA
APPLICANT: MOSKOWITZ, HAIM
APPLICANT: REISS, YUVAL
APPLICANT: ALROY, IRIS
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
FILE REFERENCE: PLV-001.01
CURRENT APPLICATION NUMBER: US/10/097,534
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/275,224
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/308,958
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/340,170
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 927
TYPE: PRT
ORGANISM: Homo sapiens
-10-097-534-15

Query Match      32.4%; Score 1259; DB 15; Length 927;
Best Local Similarity 33.6%; Pred. No. 6.1e-106;
Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

7 IRLTVLCAKWLAKDFRLLPDPKIV-----VDGSGQCHSTTVKNTLDPKN----- 55
48 VVRVVIAGIGLAKDILGASDPYRVRLTYPDPMXGLTSVQTKIKSLNPKXWBEILFRV 107
56 ---QH-----YDLVVGKTD--SITISVMN-----HKIKHKQG 83
108 HPQCHRLLEFVFDENLRTDDFLGQVDVPLYPFTENFLERPYTFKDFVLHPHSKSRV 167
84 AGFLGCVRLLSNAISRLKDT-----GYQLD-----LCKL----- 113
168 KGYL-----RLKMTYLFKTSGBDDNAEQAELEFGWVVDQPDAAACHLQQQOE 216
114 -----NPSD--TDVARGQIVVSLQ-----TEDR 134
217 PSLPLPGWEERQDILGRYYVNHESRTQWKPTPDQNLTDENGINIQACRAFTTRQ 276
135 IGTGGSVVDRCGLLEN-----EGTVYEDSG-PGRPLSCFMEEPAPYTD----- 176
277 ISBETESVDNQESSENWEIREDATMYSSQAFPPSPSSNLDVPTHLAELNARLTIFG 336
177 ----STGAAGGNCR-----FVESPS-----QDORLQAQRL-- 204
337 NSAVSQPASSNHSRRGLQAYTHEEQTLVPLLTSSGLPPGWEEKQDERGSRXYVDH 396
205 -----RNPDRGSLQTPQ-----NRPHGHSQSPPELPEGYEQ 234
397 NSRTTWTWTKPTVQATVETSQLTSSQSSAGPQSQASTSDGQQVTPSPSIEQGLPKGNEV 456
235 RTTVQGVVFLHTQGTGVSTWHDPR--IPRLNSV-----NCDLGLPLPGWVRVTSVGR 288
457 RHAPNRPFFIDHNTKTTWEDPRUKIPAHLEKGTSLDTSNDLGLPLPGWEERTHTDGR 516
289 YFVDHNNRTTQTDPRLLHMHQCLKEPSQPLPLPSGSGLEDEELPAQRYEDLDVQKL 348
517 FYINHNIKRTOWEDPRELV-----AITGPVP-----YSIDYKRY 553
349 KVLRLHLSLQOQAGHCRLEVERSEELFEESYQIWMKPKD-LKRLMKVRGREGLDYG 407
554 EFPRLKKQNDIPNKFEMKRLRATVLEDSYRRINGVRADFLKRLWIEPDGEGKLDYG 613
```

408 GYAREWLYLLCHEMLNPYYGLPOYSTDNIMYMLQINPDSSI-NPDHLSYFHFVGRIMGLAV 466
614 GVAREWFFLISKEMFNYYGLPEYSATDNYTLQINPNSGLCNEHLSYFKFGRVAGNAV 673
467 FHGYINGGFTVPFYKQILGKPIQLSDLESVDPELHKSILVWILENDITPVLDDHTFCVBEHN 526
674 YHGKLDGFFIRFFYKWLHKPITLHDWESVDSEYNSLRWILENDPTE-LDLRFIDEE 732
527 AGRILCHELKPENGVNPTSENKKEYVRLVYNWRPMRGIEAQFLAKOGFNELIPOHLL 586
733 LFGQTHQHELKNGSGSEIVVTKNKEYLYLVIOFRFVNRIOKQMAAFKEGFFELIPQDL 792
587 KPFDQKELELIIGGLDKIDLNDWKSNTRLKH-CVADSGNIVWFQAVETEDERRARLLQ 645
793 KIFDENELELLMCGLDGVDVNDREHTKYKNGYSANHQVIOFWKAVLMDSEKRIILLQ 852
646 FVTGSTRVPLQGFKAQSGTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPYSEYK 705
853 FVTGSTRVPMNGFAELYGSN---GFSQSVTEQW-GTPEKLPRAHTCFNRDLDPYSEFEE 908
706 LYEKLLTAVEETCGF 720
909 LWDKLMQMAIENTQGF 923

RESULT 9
US-10-205-823-279
Sequence 279, Application US/10205823
Publication No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Wasey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 279
LENGTH: 911
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-279

Query Match 32.2%; Score 1249.5; DB 15; Length 911;
Best Local Similarity 34.5%; Pred. No. 4.4e-105;
Matches 303; Conservative 120; Mismatches 243; Indels 211; Gaps 24;
17 LAKDFFFLP---DPPAKI---VVDGSGQCH--STDVTKNTLDPKWNQHYDLVVGKTD-S 67
69 LPTSELEFLPGICDPYVKLSLYVADENRELALVQTKTKTLAKPKNEEFYFRVNSNR 128

725 DLPPYKSYEQLEKELLPALTEETEGQCE 752

ULT 13

10-313-955-6

sequence 6, Application US/10313955

Publication NO. US20030199036A1

GENERAL INFORMATION:

APPLICANT: Beach, David H.

Caligiuri, Maureen

Nefsky, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/313,955

FILING DATE: 05-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/539,205

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 834 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

10-313-955-6

Query Match 31.6%; Score 1228.5; DB 12; Length 834;

Best Local Similarity 40.4%; Pred. No. 3.2e-103;

Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

93 LSNAISRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVSLQTRDR 134

186 LSELSRLQITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQGLPPPSVAYVHTTPG 245

135 IGTGGSV-VDCRGLL-----ENEGTVYEDSGPGRPLSCFMEEPAPYTDSTGAAGGNN-- 186

246 LPSCWEERKDAKGRTYVYVHNHNRITWT-----RPIQLAEDGA-----SGSATNSNHL 295

187 -----CRFVESPSQDQRLQALRNPDRVGLS-----QTPQNRPHGHQSPE----- 227

296 LEPQIRRRPRLSSPVTLSAPLEGAQDSVRAVKDTLSNPQSPQSPYNSPKPQKVTQ 355

228 --LPGYQRTVQGVYFLRTQGVSTWHDPRIP-----RDLNSVNCDELGLPQGWEE 280

356 SFLPPGWMRIAPNGRPFFIDHNTKTWEDPRLKFFVHMRKSTLSLNPDLGLPLPGWEE 415

281 RSTVSGRIYFVDHNRRTQFTDPRLLHIMNHQCQLKEPSPQLPLPSGSLDEELPAQY 340

416 RHLDGRFYIDHNSKIQWEDPRLQN-----PAITG-----PAPY 452

QY 341 ERDLVQKLVLRHLSLQOPQAGHCRIEVSREEIFESYRQIMK-RPKOLKKELMYKER 399

Db 453 SREFKQKYDYPRKLLKPADIPNRFEMKLRHNNIFESSYRRIMSVKPPDVLKARLWIEFE 512

QY 400 GEGLDYGGVARENLVLLCHEMLNPYYGLFOYSTDNITMQLQINPDSSI-NPDHLSYFHFV 458

Db 513 SEKGLDYGGAAREWFFLLSKEMFNYYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFI 572

QY 459 GRIMGLAVFHGHYINGGTVPFYKQLLCKPTQLSDLESVDPELHKSLSVWILENDITPVL 518

Db 573 GRVAGLAVFHGKLLDGFPIRFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 631

QY 519 HTPFCVHNAGFRILQHELKPNGRNVPTVEENKKEYRVLVYVNRVWRMGIEAQFLALQKGF 578

Db 632 LMECDEENFGQTYQVDLKPNGSEIMVTNENKREVIDLVIQWFRVNRVQKQMNAPLEGFT 691

QY 579 ELIPQHLLKPFQCKELELIIGGLDKIDLNDKNSNTRLKHCVADSN-IVRWFWQAVEFDE 637

Db 692 ELLPDLIKIPDENLELELLMCGLDVNDWRQHSIYKNGYCPNHPVQWFWKAVLLMDA 751

QY 638 ERRARLLQFVTGSTRVPLQGGFKALQGGTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDI 697

Db 752 EKRIELQLFVTGSTRVPLNGFAELYGSN---GPQLFTIEQW-GSPEKLPRATTCFNRDL 807

QY 698 PYESYEKLYEKLTAFAVETCGF 720

Db 808 PYTEFEDLRKLLMAVENAQGF 830

RESULT 14

US-10-097-534-9

; Sequence 9, Application US/10097534

; Publication No. US20030049607A1

; GENERAL INFORMATION:

; APPLICANT: GREENER, TSVIKA

; APPLICANT: MOSKOWITZ, HAIM

; APPLICANT: REISS, YUVAL

; APPLICANT: AUKOY, IRIS

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL

; TITLE OF INVENTION: MATURATION

; FILE REFERENCE: PLV-001.01

; CURRENT APPLICATION NUMBER: US/10/097,534

; PRIOR FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/275,224

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: 60/308,958

; PRIOR FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: 60/340,170

; PRIOR FILING DATE: 2001-12-07

; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 9

; LENGTH: 995

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-097-534-9

Query Match 31.6%; Score 1228.5; DB 15; Length 995;

Best Local Similarity 40.4%; Pred. No. 4.3e-103;

Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

QY 93 LSNAISRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVSLQTRDR 134

Db 347 LSELSRLQITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQGLPPPSVAYVHTTPG 406

QY 135 IGTGGSV-VDCRGLL-----ENEGTVYEDSGPGRPLSCFMEEPAPYTDSTGAAGGNN-- 186

Db 407 LPSCWEERKDAKGRTYVYVHNHNRITWT-----RPIQLAEDGA-----SGSATNSNHL 456

QY 187 -----CRFVESPSQDQRLQALRNPDRVGLS-----QTPQNRPHGHQSPE----- 227

Db 457 LEPQIRRRPRLSSPVTLSAPLEGAQDSVRAVKDTLSNPQSPQSPYNSPKPQKVTQ 516

```
228 --LPEGYQRTTVOGVFLHTQTGVSTWHDPRIP-----RDLSVNCDELGLPFGWEV 280
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 SFLPPGWNRIAPNGRPFFIDHTTTWEDPRKLPVPHMSKTSLNPNDLGLPFGWEE 576
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 RSTVSGRIYFVDHNNRTTQTPRLHHIMHQCQLKPSQPLPLPSEGSLEDELPAQRY 340
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
577 RIHLDRGTFFIDHNSKITQWEDPRLQN-----PAITG-----PAVPY 613

341 ERDLVOKLVLRHLSLQOQAGHCRIEVSREBEFESYRQIMKM-RPKDLKRLMKVFR 399
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
614 SREFKQKYDYFRKLLKPADIPNPFEMKLHNNIFESYRIRMSVKRPOVLKARLWIEFE 673
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 GEGLDYGGVAREWYLLCHEMLNPYYGLFOYSTDNIYMLQINPDSSI-NPDHLSYHFV 458
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
674 SEGGLDYGGVAREWFLLSKEMFNYYGLFEYSATDNTYTLQINPNSGLCNEDHLSYFTFI 733
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
459 GRIMGLAVFHGHIYNGGFTVPFYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLD 518
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
734 GRVAGLAVFHGKLDGFFIRFPFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 792
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
519 HTFCVHNHAFGRILQHELKPNRNVPTENKKEYRVLVYVNRWFMRGIEAOFALQGFN 578
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
793 LMFCTDEENFGQTYQVDLKPNGSBIWVTNENKREYIDLVIQWRFVNRVQKQMAFLEGFT 852
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
579 ELIPQHLKPFDOKELELIIGGLDKIDLNDWKSNTLKHCVADSN-IVRWFQAVETFE 637
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
853 ELLPIDLIKIFDENELELMCGLDGVDVNDWRQHSIYKNGYCPNHPVQWFKAVLLMDA 912
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
638 ERARLLOFVTGSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLPAKHTCFNRIDI 697
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
913 EKRIELLOFVTGSTRVPMNGFAELVGSN---GPQLFTIEQW-GSPEKLPRHTCFNRDL 968

698 PYESYEKLYEKLITAVEETCGF 720
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
969 PPYETFEDELREKLMAVENAQGF 991
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

sequence 275, Application US/10205823
Application No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Glatc, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF PROSTATE CANCER
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
FILE REFERENCE: VRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 275
LENGTH: 993
```

```
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-275

Query Match      31.6%; Score 1228.5; DB 15; Length 995;
Best Local Similarity 40.4%; Pred. No. 4.3e-103;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

QY 93 LSNAISRL-----KDTGYQRLDLCKLNPDS-----TDVARGQ-----IVVSLQTRDR 134
DB 347 LSELSRLQITPDSNGEQFSSLTQREPSSRLRSCSVTDAVAEQHLPSPSVAYVHTTGP 406
QY 135 ICTGGSV-VDCRGLL-----ENEGTVYEDSGPRPLSCFMEEPAPYTDSTGAAAGCN-- 186
DB 407 LPSGWEEKDAKAGTYVYVNNRITTTWT-----RPIQLAEDCA-----SGSATNNHNL 456
QY 187 -----CRFVESQSQRLOAQELRNPVDRGSL-----QTPQNRPHCHQSP----- 227
DB 457 IEPQIRRPRLSSPTVTLSPLEGAKDSPVRAVKDTLSNPQSPQSPSPYNSPKPQHKVTQ 516
QY 228 --LPEGYQRTTVOGVFLHTQTGVSTWHDPRIP-----RDLSVNCDELGLPFGWEV 280
DB 517 SFLPPGWNRIAPNGRPFFIDHTTTWEDPRKLPVPHMSKTSLNPNDLGLPFGWEE 576
QY 281 RSTVSGRIYFVDHNNRTTQTPRLHHIMHQCQLKPSQPLPLPSEGSLEDELPAQRY 340
DB 577 RIHLDRGTFFIDHNSKITQWEDPRLQN-----PAITG-----PAVPY 613
QY 341 ERDLVOKLVLRHLSLQOQAGHCRIEVSREBEFESYRQIMKM-RPKDLKRLMKVFR 399
DB 614 SREFKQKYDYFRKLLKPADIPNPFEMKLHNNIFESYRIRMSVKRPOVLKARLWIEFE 673
QY 400 GEGLDYGGVAREWYLLCHEMLNPYYGLFOYSTDNIYMLQINPDSSI-NPDHLSYHFV 458
DB 674 SEGGLDYGGVAREWFLLSKEMFNYYGLFEYSATDNTYTLQINPNSGLCNEDHLSYFTFI 733
QY 459 GRIMGLAVFHGHIYNGGFTVPFYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLD 518
DB 734 GRVAGLAVFHGKLDGFFIRFPFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 792
QY 519 HTFCVHNHAFGRILQHELKPNRNVPTENKKEYRVLVYVNRWFMRGIEAOFALQGFN 578
DB 793 LMFCTDEENFGQTYQVDLKPNGSBIWVTNENKREYIDLVIQWRFVNRVQKQMAFLEGFT 852
QY 579 ELIPQHLKPFDOKELELIIGGLDKIDLNDWKSNTLKHCVADSN-IVRWFQAVETFE 637
DB 853 ELLPIDLIKIFDENELELMCGLDGVDVNDWRQHSIYKNGYCPNHPVQWFKAVLLMDA 912
QY 638 ERARLLOFVTGSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLPAKHTCFNRIDI 697
DB 913 EKRIELLOFVTGSTRVPMNGFAELVGSN---GPQLFTIEQW-GSPEKLPRHTCFNRDL 968
QY 698 PYESYEKLYEKLITAVEETCGF 720
DB 969 PPYETFEDELREKLMAVENAQGF 991
```

Search completed: February 20, 2004, 15:37:43
Job time : 35.9477 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

n on: February 20, 2004, 15:25:26 ; Search time 17.2026 Seconds
(without alignments)
4041.832 Million cell updates/sec

file: US-10-009-945-2

rfect score: 3884
xact: 1 GSSSKIRLTVLCANLAKK.....EKLVEKLLTAVETCGPAVE 723

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 283308 seqs, 96168682 residues

tal number of hits satisfying chosen parameters: 283308

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

tabase : PIR.76:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1492.5	38.4	767	2 T37545	ubiquitin-protein
2	1471	37.9	766	1 S66562	ubiquitin-protein
3	1437	37.0	815	2 T49744	probable ubiquitin
4	1377	35.5	809	1 S43217	ubiquitin-protein
5	1341	34.5	786	2 T39585	ubiquitin-protein
6	1241.5	32.0	820	2 T46412	ubiquitin-protein
7	1227.5	31.6	887	2 S70642	ubiquitin ligase N
8	1051.5	27.1	708	2 I83196	NEEDD-4 ORF - mouse
9	1026.5	26.4	671	2 T37900	probable ubiquitin
10	854.5	22.0	3839	2 T49799	related to TOM1 pr
11	848	21.8	3227	2 T37964	protein F147L6.10
12	784	20.2	4056	2 H36599	ubiquitin-protein
13	780.5	20.1	1126	2 T01421	ubiquitin-protein
14	769	19.8	3268	2 S69625	hypothetical prote
15	669	17.2	889	2 T20274	hypothetical prote
16	627.5	16.2	310	2 I52645	DNA binding protei
17	604.5	15.6	221	2 T51886	hypothetical prote
18	530.5	13.7	874	2 A38920	EG-associated prot
19	484	12.5	807	2 T40821	probable ubiquitin
20	481.5	12.4	892	2 S70355	probable membrane
21	471	12.1	1142	2 T46155	hypothetical prote
22	467	12.0	1083	2 A38919	hypothetical prote
23	466	12.0	873	2 H85134	polyubiquitin-like
24	451	11.6	1029	2 T38951	probable ubiquitin
25	449.5	11.6	1054	2 B38919	hypothetical prote
26	444	11.4	1066	2 T36753	hypothetical prote
27	418	10.8	1001	2 T20373	hypothetical prote
28	400.5	10.3	655	2 T41750	probable ubiquitin
29	382.5	9.8	4836	2 T14346	herc2 protein - mo

ALIGNMENTS

RESULT 1

T37545
ubiquitin-protein ligase (EC 6.3.2.19) pub1 [validated] - fission yeast (Schizosacchar
N; Alternate names: ubiquitin ligase Pub1
C; Species: Schizosaccharomyces pombe
C; Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Jun-2002
C; Accession: T37545; T48655
R; McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
Submitted to the EMBL Data Library, September 1997
A; Reference number: Z21722
A; Accession: T37545
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: DNA
A; Residues: 1-767 <MCL>
A; Cross-references: EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GN00066; SPDB:SPAC11G7.02
A; Experimental source: strain 972h-; cosmid c11G7
R; Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.
Mol. Gen. Genet. 254, 520-528, 1997
A; Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning pub1
A; Reference number: Z07985; MUID:97340937; PMID:9197411
A; Accession: T48655
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: DNA
A; Residues: 1-767 <SAL>
A; Cross-references: EMBL:U62795; NID:G2262192; PIDN:AA863350.1; PID:G2262193
A; Experimental source: strain J227
C; Genetics:
A; Gene: pub1; SPDB:SPAC11G7.02
A; Map position: 1
A; Introns: 6/2: 14/1, 62/2
C; Function: <UBI>
A; Description: EC 6.3.2.19 [validated, MUID:96205868]
C; Function: <CYC>
A; Description: involved in of the mitotic activating tyrosine phosphatase cdc25 [valid
C; Function: <TOL>
A; Description: required for low pH-tolerance [validated, MUID:97340937]
C; Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C; Keywords: cell cycle control; ligase
F; 205-242/Domain: WW repeat homology <WW1>
F; 288-325/Domain: WW repeat homology <WW2>
F; 345-382/Domain: WW repeat homology <WW3>

Query Match 38.4%; Score 1492.5; DB 2; Length 767;

Best Local Similarity 41.1%; Pred. No. 1.7e-100;

Matches 329; Conservative 119; Mismatches 234; Indels 119; Gaps 19;

QY 4 SIKRLTVLCANLAKKDFRLDPFAKIVDSCGQCHSDTVKNTLDPKWHQVDLYVG 63

DB 7 SRRIRVIVAAADGLYKEDVRFDPFAVLTVDSGE-QTHHTTAIKTLNFWNETFEVNT 65

QY 64 KTDSTITVWNHKKHKQAGFLGCVRLLSNAISRLKDTGYQRL--DLCKLNPSTDAV 121

```

66  DNSTIAIQVFDQKF-KKGGQGLGVINLRVGVLDLAIGGDEMLTRDLKSN--ENTVV 122
122  RGQIVVSLQ-----TRDRIGGSDVDCRGLLENEGVYED 157
123  HGKLIINLSTAAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTRD 182
158  SGP-----RPLSCFMBE---PAPY---TDSTG-----AAA 182
183  NAPAASPASSERPTFSFEDQYGLPFGWEERTDNLGRYYVDHNRSTTWRPNLSSVA 242
183  CGGNCRFVESPS-----QDRLQARLNPDVRSLOTPQNRPHGHCSELPEGVYED 234
243  GAAAEHLSSASSANVTGVPSSSNAARTEASVLTSNATTAG-----SGELPPGHEQ 296
235  RTVVGQVYFLHTQGVSTWDPRIPDLNSVN-----CDELGPLPGWEVRS 283
297  RYTPGGRFYFDHNRTRTTTWDPRQYIRSYGGPNNAITQQQPVSQLGFLPSGWEMLT 356
284  VSGRIYFVDHNRTRTTQDTPRLHHIMNHQCOLKEPSOPLPLPSEGSLDEELPAQRYED 343
357  NTARYVYFDHNRTRTTTWDPR-----LPS--SL--DQNVF--QYKED 392
344  LVOKLVLRLHLSLQQAQCHRIEVSREEIFEESYRQIMKMRPKDKRLMVKPRGEG 403
393  FRKLIYLSQPAL-HPLPGQCHIKVRNHI-FEDSAEIMRQSATDLKRLMIKFDGEDG 451
404  LDYGGVAREWLYLLCHEMLNPPYGLFOYSTDNIMQLQINPDSSINPDHLSYFHFVGRIMG 463
452  LDYGLSREYFLLSHEMFNFYCLFEYSVDNYTLQINPHSGINPEHLNFKFGRVIG 511
464  LAVFHGHVINGGFTVPYKOLLGKPIQLSDLESVDPELHKLVLWILENDITVLDHTFCV 523
512  LAIFHRRFVDAFFVVSFYKMLQKVTLOQMESDAEYRSLWILDNDITGVLDLTFV 571
524  EHNARGLRIQHELKPNRNPVTEENKKEVYRLYVNRWFRMGTEAQFLAQKGFNELIPQ 583
572  EDCNCGEVVITDLKPNRNIETVEENKKEVYRLYVNRWFRMGTEAQFLAQKGFNELIPQ 631
584  HLKPFDOKELELIGLDKIDLNDKSNRLKHCVDASNIVRWFQAVETFEERRARL 643
632  ELINVFDERELELLIGLSEIDMEDWKHKDYRSYSENQDIQKFWELMDEWSNKKSL 691
644  LQFVTGSTRVPLQGFKAQGSTGAAGPRLETHILIDANTDNLKHAHTCFNRLDIPYTSK 703
692  LQFTTGTSRIFVNGFKDLQSGD---GPRKFTIEKA-GEPNKLPKHAHTCFNRLDLPYTSK 747
704  EKLVEKLTAVEETCGFAVE 723
748  KOLDHKLSTIAVEETIGFQGE 767

```

RESULT 2

```

ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)
;Alternate names: E6-AP-like protein ubiquitin ligase
;Species: Schizosaccharomyces pombe
;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
;Accession: S66562; T45159
;RefSeq: B.; Beach, D.
;Title: Pubi acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25
;Reference number: S66562; MUID:96205868; PMID:8635463
;Accession: S66562
;Status: preliminary; nucleic acid sequence not shown
;Molecule type: mRNA
;Residues: 1-766 <NER>
;Cross-references: GB:U66716; NID:g1519443; PID:AAS07514.1; PID:g1519444
;RefSeq: B.S.; Beach, D.
;Submitted to the EMBL Data Library, August 1996
;Description: Pubi acts as an E6-AP-like protein ubiquitin ligase in the degradation of
;Reference number: Z22935
;Accession: T45159
;Status: preliminary; translated from GB/EMBL/DBD

```

```

A:Molecule type: mRNA
A:Residues: 1-766 <NER>
A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1
C:Genetics:
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C:Keywords: ligase
F:205-242/Domain: WW repeat homology <WW1>
F:288-325/Domain: WW repeat homology <WW2>
F:345-382/Domain: WW repeat homology <WW3>

Query Match 37.9%; Score 1471; DB 1; Length 766;
Best Local Similarity 40.9%; Pred. No. 6.3e-95; Indels 120; Gaps 20;
Matches 327; Conservative 118; Mismatches 235;

QY 4 SIKRLTVLCARNLAKKDFRLLPDPFAKIVVDSGGCHSTDTYKNTLDPKWNQHYDLYVG 63
DB 7 SRIRVTVAADGLYKRDVFRFPDPFAVLTVDGE-QTHTTTAKKTLNPNWNETPEVNT 65
QY 64 KTDSTISVMNHKHKQAGFLGCVRLLSNLSLTKDTGYQL--DLCKLNPSTDVA 121
DB 66 DNSTIAIQVFDQKF-KKGGQGLGVINLRVGVLDLAIGGDEMLTRDLKSN--ENTVV 122
QY 122 RGQIVVSLQ-----TRDRIGGSDVDCRGLLENEGVYED 157
DB 123 HGKLIINLSTAAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTRD 182
QY 158 SGP-----RPLSCFMBE---PAPY---TDSTG-----AAA 182
DB 183 NAPAASPASSERPTFSFEDQYGLPFGWEERTDNLGRYYVDHNRSTTWRPNLSSVA 242
QY 183 CGGNCRFVESPS-----QDRLQARLNPDVRSLOTPQNRPHGHCSELPEGVYED 234
DB 243 GAAAEHLSSASSANVTGVPSSSNAARTEASVLTSNATTAG-----SGELPPGHEQ 296
QY 235 RTVVGQVYFLHTQGVSTWDPRIPDLNSVN-----CDELGPLPGWEVRS 283
DB 297 RYTPGGRFYFDHNRTRTTTWDPRQYIRSYGGPNNAITQQQPVSQLGFLPSGWEMLT 356
QY 284 VSGRIYFVDHNRTRTTQDTPRLHHIMNHQCOLKEPSOPLPLPSEGSLDEELPAQRYED 343
DB 357 NTARYVYFDHNRTRTTTWDPR-----LPS--SL--DQNVF--QYKED 392
QY 344 LVOKLVLRLHLSLQQAQCHRIEVSREEIFEESYRQIMKMRPKDKRLMVKPRGEG 403
DB 393 FRKLIYLSQPAL-HPLPGQCHIKVRNHI-FEDSAEIMRQSATDLKRLMIKFDGEDG 451
QY 404 LDYGGVAREWLYLLCHEMLNPPYGLFOYSTDNIMQLQINPDSSINPDHLSYFHFVGRIMG 463
DB 452 LDYGLSREYFLLSHEMFNFYCLFEYSVDNYTLQINPHSGINPEHLNFKFGRVIG 511
QY 464 LAVFHGHVINGGFTVPYKOLLGKPIQLSDLESVDPELHKLVLWILENDITVLDHTFCV 523
DB 512 LAIFHRRFVDAFFVVSFYKMLQKVTLOQMESDAEYRSLWILDNDITGVLDLTFV 571
QY 524 EHNARGLRIQHELKPNRNPVTEENKKEVYRLYVNRWFRMGTEAQFLAQKGFNELIPQ 583
DB 572 EDCNCGEVVITDLKPNRNIETVEENKKEVYRLYVNRWFRMGTEAQFLAQKGFNELIPQ 630
QY 584 HLKPFDOKELELIGLDKIDLNDKSNRLKHCVDASNIVRWFQAVETFEERRARL 643
DB 631 ELINVFDERELELLIGLSEIDMEDWKHKDYRSYSENQDIQKFWELMDEWSNKKSL 690
QY 644 LQFVTGSTRVPLQGFKAQGSTGAAGPRLETHILIDANTDNLKHAHTCFNRLDIPYTSK 703
DB 691 LQFTTGTSRIFVNGFKDLQSGD---GPRKFTIEKA-GEPNKLPKHAHTCFNRLDLPYTSK 746
QY 704 EKLVEKLTAVEETCGFAVE 723
DB 747 KOLDHKLSTIAVEETIGFQGE 766

```

RESULT 3

9744 obable ubiquitin-protein ligase [imported] - Neurospora crassa
Alternate names: protein B24B19.160
Species: Neurospora crassa
Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
Accession: T49744
Schulte, U.; Aign, V.; Kohseisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, B.
Submitted to the Protein Sequence Database, May 2000
Reference number: Z25022
Status: preliminary
Molecule type: DNA
Residues: 1-815 <SCH>
Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.160
Experimental source: BAC clone B24B19, strain OR74a
Genetics:
Gene: NCSP:B24B19.160
Map position: 6
Introns: 11/1; 24/1; 59/2; 110/1; 783/2
Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
219-276/Domain: WW repeat homology <WW1>
334-371/Domain: WW repeat homology <WW2>
393-430/Domain: WW repeat homology <WW3>
Query Match 37.0%; Score 1437; DB 2; Length 815;
Best Local Similarity 38.3%; Pred. No. 2.1e-96;
Matches 321; Conservative 125; Mismatches 238; Indels 154; Gaps 21;
9 LTVLCAKNIKKAKDFRLLPDPFAKIVDGGSGCHSDTVKNTLDPKWNHGYDLV-----V 62
9 LPVIAADGLYKRVDFRPPFPFVATNGE-QTKTVQSKRTLPNFWNEHDFWLIQSKV 67
63 GKTDSTITISVNNHKKHKQAGFLGCVRL-LSNAISRLKDTGYQRLDCLKNPDSPTDAV 121
68 NEDSILAVQVFDQKF-KKXQDQGLGVINVRIGDVIELAPDAEDQMLRDLKSKDNLV 126
122 RQIVSVLQTR-----DRITGSGVDCRGLLENCTVYED-----SGP----- 160
127 HGKLIINLSNLTATMSRLGFPSPSSRLTLTQSSVIGSNDRANERPSSMSGNGTANN 186
161 ----GRPLSCFMEE-----PA-----PYTDSGT----- 179
187 MTLASRPASLAVSSSTAPPTGNGTAPNPSLTVPAQARHSHSTLSPEDSNGRLPAGWE 246
180 -----AAAGGNCRFVSPSQDORLOAQRINPDVRSQLOT 215
247 RREDHLGRTYVDHNSRTTSMNPTGTGAENRTABANTQVERQHRNLTLPEDRTGANS 306
216 PQNRPH-----GHOSP---ELPEGYEORTVQGVYFLHTQTGVSTWHDPR 258
307 PTLOQQQAATANATMMHTGATPTGELPAGWEQRTPEGEPYFVDHNTTITTTWDP 366
259 IPRDL-----NSVN-----CDELGLPDPGWEVSTVSGRIYFVDHNNRTQTDPRL 305
367 RQVIRMYGQNTNGTITQQQVSQLGPLSPGWEMLTNTARVYFVDHNTKTITWDDPR- 425
306 HHINWQCQLKEPQPLPSEGLSEDEELPAQYERDLVKLVLRHLSLQQQPAQHC 365
426 -----LPS--SL--DQNVF---QYKDRFRLKLIIFRSOPAM-RWMSQC 461
366 RIEVSREEIEESYRQIMKRPKDLKKRLMKVPRGEGLDYGVAREWLYLCHENLNPY 425
462 HIKVRRSHIFEDSPAIESRQSAATDKRLMKIKDGEDGLDYGSLSEFPFLLSHEMFNP 521
426 YGLQYSTDNTYMLQINPDSINPDHLSYHFHVRINGLAVPHGVINGFTVYFKQL 485
522 YCLFEYSAHDNTYIQINPHSGINPHLNYPKFIRVVVGIAIFRRFLDAFFGALYKWL 581
486 GKPIQLSDLESVDPELHKSLVILENDITVLDHTFCVEHNAFRQILOHELKPNGRNVFV 545
582 GKAVSLADMEGVADAFHRSQWMLDNDITDVLDTATSTEDERGVITEEDLIPNGRIAV 641
546 TEENKKEVRLVNVWRFWRGIEAQFLAQGFNELIPQHLKPFQDQKELELLIGLDKID 605

Db 642 TNEKKEVELMVAWREKLEQQFRAKDFGFLHLPQDLINVFDERELLELLGGIAEID 701
QY 606 LNDWKSNTRLKXCHVADSNVWFWQAVETDEERRARLLQFVTGSTRVLPQKALQGST 665
Db 702 VDDMKKHTDYRGYTESDEVTFQFQVTVRSDWGEQKSLQFOTTTGTRIPVNGPKDJQSGD 761
QY 666 GAAGRLTLTHLIDANTDNLPAKATCFNRIDIPPEYSEKLYEKLTAVEETCGFAVE 723
Db 762 ---GPRRTTIEKAGEIT-NLEKATCFNRDLDPFKSLEMLQKLLTAVEETWGFQGE 815
RESULT 4
S43217
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)
N/Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein
C/Species: Saccharomyces cerevisiae
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C/Accession: S43217; S50628; S70050
R/Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor
submitted to the EMBL Data Library, February 1993
A/Reference number: S30812
A/Accession: S43217
A/Molecule type: DNA
A/Residues: 1-809 <MUL>
A/Cross-references: GB:U18916; EMBL:L11119; NID:G1384128; PID:AC03223.1; PID:G603364
R/Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A/Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda
A/Reference number: S50628
A/Accession: S50628
A/Molecule type: DNA
A/Residues: 1-809 <DIE>
A/Cross-references: EMBL:U18916; NID:G1384128; PID:AC03223.1; PID:G603364; GSPDB:GN0
Mol. Microbiol. 18, 77-87, 1995
A/Title: NPI1, an essential yeast gene involved in induced degradation of Gap1 and Fur
A/Reference number: S70050; MUID:96154942; PMID:8596462
A/Accession: S70050
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-101 <HEI>
C/Genetics:
A/Gene: SGD:RSP5; PUB1; NPI1; MIPS:YER125w
A/Cross-references: MIPS:YER125w; SGD:S0000927
A/Map position: 5R
C/Function:
A/Description: Involved in endocytosis of GAP1 protein and FUR4 protein; binds and ubi
C/Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C/Keywords: ligase
F/229-266/Domain: WW repeat homology <WW1>
F/331-368/Domain: WW repeat homology <WW2>
F/387-424/Domain: WW repeat homology <WW3>
Query Match 35.5%; Score 1377; DB 1; Length 809;
Best Local Similarity 37.0%; Pred. No. 4.8e-92;
Matches 313; Conservative 131; Mismatches 238; Indels 164; Gaps 24;
QY 3 SSIKILTVLCAKNIKKAKDFRLLPDPFAKIVDGGSGCHSDTVKNTLDPKWNHGYDL-Y 61
Db 3 SSISVXL--VAESLYKRVDFRPPFPFVATIDGY-QTKSTAAKTLNPNYNETFKFDD 59
QY 62 VGKTDSTITISVNNHKKHKQAGFLGCVRL-LSNAISRLKDT-----TGYOR-----LDL 110
Db 60 INENSILITIOVFDQKF-KKKDQGLGVNVVRVGVLDGLDEDTATSSSRPREETITRDL 118
QY 111 CKLNPSDTPAVGQIVV-----SLQTRDRIGTGVSVVDCRGLL 148
Db 119 KKNK--DGMVSGRLIVLSKLPSSPSHQAQSGHTASSNTSTSTFTNGHSTSS---T 173
QY 149 ENEGTVYEDSGGRPLSCFMEEFAP-----YTDSTGAAA----- 182
Db 174 RNHSTSPSRGTAQAVESTLQSGTTAATNTATTSRSTNSTSSATROYSPEDQYGRLLP 233

```
183 -----GGGNCRFYE-----SPQDORLOAQRINPDRVGRSLQTPQNR-----PHG 222
234 GWERTDNFGRYYVDHNTRTTTWKRPDLQ-TEAERGNQLNANTELERQHRGRTLPQG 292
223 HOS-----PELPEGEYQRTTVOGQVFLHT 247
293 SDNSSVTVQGGGNNIPVNGAAAAAFAATGGTSGLGELPSSGWEORFTPEGRAVVDH 352
248 QTCVSTWHDPRIPRDLNSYN-----CDELGLPPGWESTVSGRIYFVDHNNRT 297
353 NRTTITWVDRPQQVIRTYPTNTTIQQPVSQLGPLSGHEMLTNTARVYFVDHNTKT 412
298 TQPTDRLHHIWMHQCOLKEPQOPPLPSEGSLEDEELPAQRYERDLVQKLVLRHELSL 357
413 TTWDDPR-----LPS--SL-DQNVF--QYKRDPRKVIYFRSPAL 448
358 QPQAGHCRIEVSREIEFESYRQIMKRPDKLKKLMVKFRGEGLDYGGVAREMIYLL 417
449 -RLPGQCHIKVRKNIPEDAYQEIQRQTFEDLKKRLMIKFDGEGLDYGGVAREFELL 507
418 CHEMLNPIYGLFOYSTDNITMQLNPDSINPDHLSYFHFVGRINGLAHVGHYINGFT 477
508 SHMPNPFYCLPEYSAYDNYTQINPNSGINPEHLNYPKFGVGVGLVFRHRLDAFFV 567
478 VPYKOLLGPKIOLDSLESVDPELHKSILVILENDITPVLDHTFCVEHNAFRGILQHELK 537
568 GAIYKMKLKKVVLQDVEGDAEVVNSLWMLNSIDGVLDLTFESADDERFGEVTVDLK 627
538 PNGRNVPVTEENKEVRYLVNWRFRMGTEAQLQKGFNELIPQHLLKFPDQKELELI 597
628 PDGRNTEVTDGNKEVYELQWRIYDVRQEQKAFMDGFGNELIPEDLVTVFDERELELL 687
598 IGGDLKIDLNDKSNTRLKHCVADSNIVWFQAVETFEERRARLLQFVTGSRVPLQG 657
688 IGGIAEIDEDKXNDYRQYQSDDEVIQFWKVCSEWNEQEARLLQFTGTSRIPVNG 747
658 FKALQSTGAGPRLFTIHLIDANTNLNPKAHTCFNRIDIPPEYSEYKLEKLTAVEET 717
748 FKDLQSGSD--GFRFTIEKA-GEVQQLPKSHTCFNRVLDLPQYVDYDSMKOKLTLAVEET 803
718 CGFAVE 723
804 IGFQGE 809
RESULT 5
ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)
Species: Schizosaccharomyces pombe
Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
Accession: T39585
Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, August 1997
Reference number: 221865
Accession: T39585
Status: Preliminary; translated from GE/EMBL/DBJ
Molecule type: DNA
Residues: 1-786 <VOL>
Cross-references: EMBL:299759; PIDN:CAH16903.1; GSPDB:GN00067; SPDB:SPBCL6E9.11c
Experimental source: strain 972h-; cosmid c16E9
Genetics:
Gene: SPDB:SPBCL6E9.11c
Map position: 2
Introns: 60/2
Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
236-273/Domain: WW repeat homology <WWR1>
306-343/Domain: WW repeat homology <WWR2>
364-401/Domain: WW repeat homology <WWR3>
Query Match 34.5%; Score 1341; DB 2; Length 786;
Best Local Similarity 36.6%; Pred. No. 1.9e-89;
Matches 304; Conservative 128; Mismatches 236; Indels 162; Gaps 21;
```

```
QY 6 KIRLTVLCAKNLAKKOFERLPDPFAKIVVDGSGQCHSTDTVKNTLDPKNOHOHYDLYVGKT 65
DB 7 RVRVYVAAAGLKRDLRQPDFAITVDGB-QHTTKVKKSNPNYWNNEGFVTVKPS 65
QY 66 DSITISVNNHKKIKKOGAGFLGCVRLLSNAISRLKDTGYQR-----LDLCK-LNPSDDT- 119
DB 66 SVISIRLFDQKF-KKXDOGLGLVGF-----RMEVGSFRSNREVSRLPLKKSSTTN 118
QY 120 -AVGQIVV-----SL 129
DB 119 LSVLGNLVKVPKIRAPAGNHSSTANTTGTPTTTTARTTRTPRTATNTSNQST 178
QY 130 QTRDRIGTGGSVVDCRGLLENEGTVYEDSGFGRPL-----SCFMEBPA-----PYTOS 177
DB 179 SNSITNGTSAATNGTGTGAGTGASHRSS-----PVNTRQNTNTSALSNSNAHIMSSFEDQ 234
QY 178 TGAAGG-----GNCRFYE-----SPQDORLOAQRINPDRV 209
DB 235 YGRLLPGWREADSLGRYYVDHNTRTTTWTRPASSTNPVHNTSSDSORLNQHRPLD- 293
QY 210 RGSLLQTPONRPHGHQSP---ELPEGEYQRTTVOGQVFLHTQTCVSTWHDPRIP- 260
DB 294 -----DSNLSLWQSDSGNDLPGWEMRYTDGRPYFVDHNTTITTTWVDRNPLVRPNG 346
QY 261 -----ROLNSVNDDELGLPLPPGWESTVSGRIYFVDHNNRTQPTDRLHHIWMHQCO 314
DB 347 GSSITVGSMLQPSLSHLGLPLSGWEMELTNSARVYFVDHNTKTITWDDPR- 397
QY 315 LKPEQSLPLPSEGSLEDEELPAQRYERDLVQKLVLRHELSLQOPQAGHCRIEVSREI 374
DB 398 -----PSALDQDVP--QYKCDPRKLIYFRSQPGM-RPLPGQCNVKKVRCHI 441
QY 375 FEESYRQIMKRPDKLKKLMVKFRGEGLDYGGVAREMIYLLCHEMLNPIYGLFOYSTD 434
DB 442 FDSYAEIMRYSAHDLKKRLMIRFDGSDGLDYGSLREFFLLSHKMFDFICYLFEYSAV 501
QY 435 NYMLQINPDSSINPDHLSYFHFVGRINGLAHVGHYINGFTVPFYKOLLGPKIQLSDL 494
DB 502 DNYTQINPHSSINPEHLNYPFRFGVIGLALFRRFLDAFFVYVSLYKLLRKKVSLADM 561
QY 495 ESDVPELHKSILVILENDITPVLDHTFCVEHNAFRGILQHELKPNGRNVPVTEENKEV 554
DB 562 ESIDAEFYSLKWLNDITGILDITFSVEBDHFGVTRVETLITNGENIEVTEENKKYV 621
QY 555 RLYVNRFRMGTEAQLQKGFNELIPQHLLKFPDQKELELIIGGLDKIDLNDKSNTR 614
DB 622 DLVTEWRVSKRVEQGFNAPYSGFVELSPDLVNVFDERELELLIGGISDVDEWKSHTE 681
QY 615 LKHCVADSNIVWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFT 674
DB 682 YRYIATDPVKKWFELIAGWKNEDESKLLQFATGTSRIPVNGFRDLQSGD--GPRKFT 738
QY 675 IHLIDANT-DNLNPKAHTCFNRIDIPPEYSEYKLEKLTAVEETCGFAVE 723
DB 739 IE--KAGTPDQPLFVAHTCFNRDLDPYPSKDTLHEKSLAVENTVGFQNE 786
RESULT 6
T46412
ubiquitin-protein ligase (EC 6.3.2.19) NEDD4 - human (fragment)
N:Alternate names: hypothetical protein DKEZp4342422.1
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
C:Accession: T46412
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223034
A:Accession: T46412
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-820 <AAA>
A:Cross-references: EMBL:AL137469
```

Experimental source: adult testis; clone DKFZp34P2422

Genetics:

Gene: GDB:NEDD4

Note: DKFZp34P2422.1

Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

Keywords: ligase

142-179/Domain: WW repeat homology <WWR1>

342-379/Domain: WW repeat homology <WWR2>

393-430/Domain: WW repeat homology <WWR3>

489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 32.0%; Score 1241.5; DB 2; Length 820;

Best Local Similarity 34.9%; Pred. No. 3.7e-82;

Matches 294; Conservative 116; Mismatches 230; Indels 203; Gaps 21;

43 TDVNTLTDPKQNHVDLYVGKTD-SITTSVWNKKHKQAGFLGCVRLLSNAI----- 97

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 97

12 TKTIKTLNPKWNEEFVFNPSNHLLEFVDENLTRDD---FLGQVDVPLSHLPTD 68

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 68

98 -----SRLDGTGYORLDLCKLNPSDSDAVRGOIVSLQTRDRTG 137

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 137

69 PTERPAPKDFLLRPRSHKSRVK--GFLRLKWAYMPKNG-----GQDENSDQDDMEH 121

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121

138 GGSVYDCRG-----LLENQGTVY----- 155

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 155

122 GWEVVDNDSASQHEBELPPPLPGWEEKVDNLGRYYVNNHNRRTQWHRPRLMDVSSE 181

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181

156 -----EISGP-----GRPLSCFWEHPAPYTDSTGAA-- 181

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181

182 SDNNIRIQNEAHRFRSRHISELEPEPSGGVDPWPETISEEVNIAGDSGLALP 241

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241

182 ---AGGNCRFVSPSQD-----QRLQAQRLN----- 206

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 206

242 PPPASPGSRSTSPQELSELSRLRLQITPDSNGEFSLLIQEPSSRLRSCSVTDAVEQGH 301

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301

207 -----PDVRSGLQTPQNRPHQSPQE-----LPEYEQRTTVQGVYEL 245

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 245

302 LPPGAKDSPVRAVKOTLSLPQSPQSPYNSPKQHKVTQSLPFPGWENRIAPNRPFFI 361

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 361

246 HTQTGVSTWHDPRIP-----RDLNSVNCDELGLPPLPGWEVRSVSGRIYFVDHNRRTQF 300

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300

362 DHNTKTTTWEDPLRLKFFVHMRKSTSLNPNLGLPLPGWEERHLDGRFTYDHSNKITQW 421

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 421

301 TDPLRLHMHNCOLKEPSOPLPLPSEGSLEBELPAQRYERDLVKLVLRHLSLQOP 360

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360

422 EDPLQN-----PAITG-----PAVPYSREFKQKYDFRKLKPKAD 458

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 458

361 QAGHCRIVSRREEIFERSYQIMKM-RPKDLKKLMVKFRGEGLDYGGVAREWLYLLCH 419

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 419

459 IPNRFEMKLRHNNIFESYKIRMSVKRPDVVKARLWIEFESEKGLDYGGVAREWFLLSK 518

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 518

420 EMLNPPYGLFOYSTDNIMQLNPDSSI-NPDHLSYFHFVGRIMGLAVFHGHYINGFTV 478

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 478

519 EMFNPPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFTGRVAGLAVFHGKLLDGFTR 578

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 578

479 PPKYKLLGKPLQLSDLESVDLHKSLLVILENDITVPLDHTFCVSHNFAFRLLQHELKPE 538

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 538

579 PPKYKLLGKPLQLSDLESVDLHKSLLVILENDITVPLDHTFCVSHNFAFRLLQHELKPE 637

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 637

539 NGRNPVPTVEENKKEYVRLVNNRMFRGIEAQFLALQGFNELIPQHLKPFDDQKELELII 598

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 598

638 NGSEIMVTNENKREYIDLVIQWRFNVRVQKQNAFLEGFTTELLPDLIKIFDENELELLM 697

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 697

599 GLDKIDLNDKSNTRLKHCVADSN-IVRPMQAVETFDERRARLLQFTVGSSTRYPLOG 657

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 657

698 CGLGDVDVNDNRQHSIYKNGYCPNHPVIQWPKAVLLMDAEKIRLLQFVIGTSRYPMNG 757

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 757

658 FYALQCGTGAAGPRFLTTHLIDANTDLNPKAHTCFNRIDIPPPVESYEKLYELLTAVEET 717

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 717

758 FAELYGSN---GPQLFTIEQW-GSPEKLFRACHTCFNLDLPPYETTFEDREKLMAVENA 813

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 813

718 CGF 720

|||||

Db 814 QGF 816

RESULT 7

S70642

ubiquitin ligase Nedd4 - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000

C;Accession: S70642

C;Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.

EMBO J. 15, 2371-2380, 1996

A;Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+

A;Reference number: S70642; MUID:96221297; PMID:8665844

A;Accession: S70642

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-887 <STA>

A;Cross-references: EMBL:U50842; NID:gl293646; PIDN:AAB48949.1; PID:gl293647

C;Genetics:

A;Gene: Nedd4

C;Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubi

F;54-157/Domain: protein kinase C C2 region homology <XC2>

F;246-283/Domain: WW repeat homology <WW1>

F;402-439/Domain: WW repeat homology <WW2>

F;459-496/Domain: WW repeat homology <WW3>

F;555-881/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 31.6%; Score 1227.5; DB 2; Length 887;

Best Local Similarity 33.9%; Pred. No. 4.4e-81;

Matches 296; Conservative 127; Mismatches 224; Indels 227; Gaps 28;

Qy 7 IRLTVLCANLAKKQFFRLDPDFAKIV---VDGSGQCHSTDTVKNLDPKWN----- 55

Db 77 VRKYVIAGIGLAKKQILGASDFYVRVTLVDPMSGVLTSTVQTKIKSLNPKWNEILFRV 136

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 136

Qy 56 ---QH-----YDLVVGKTD--SITISVWN-----HKIKHKQG 83

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 83

137 LPQHRILFEVDENLRITDDFLGQVDVPLPLPTENRMERPYTFKDFVLRPSKSRV 196

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 196

Qy 84 AGFLGCVRLLSNAIRLMDT----- 103

Db 197 KGVL-----RLKMTYLPKNGSDDENADQAELEPGWVVLDDPDAATHLQHPPEP 245

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 245

Qy 104 -----GY-QRLDLC-----KLNPSD--TDAVGOIVV-----SLQTRDRI 135

Db 246 SPLPPGWEERQDVLGRYTYVNNHESRTQWRKSPEDDITDDENGDIQLQAHGAFTRRQI 305

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 305

Qy 136 GTGGSVVDCRG-----LLENQGTVYEDSGQRPLS-----C 166

Db 306 SE-----DVDGPDNHESPENWEIVREDENTIIYSGQAVQSPSPSGHPDVQVRLAEELDTLT 360

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360

Qy 167 FMEEPA---PYTDSGTAAAGGN--CRFVESPSQORLQAQRLNPDVGRSLOTQPNRPH 221

Db 361 MYGNPATQSVTSSNHSRSGSSQTCIFEEQTLPLVLLFT----- 400

Qy 222 GHOSPLPEGEYQRTTVQGVYFLHTQTGVSTWHDPE-----RIPRDLNS---VNCD 270

Db 401 ---SSGLPPGWEEKQDGRGSIYVDHNSKTTTWSKPTMQDDPRSKIPAHLRGKTPEVDSND 457

Qy 271 LGPLPPGWVRSVSGRIYFVDHNRRTQFTPRLHMHNCOLKEPSQPLPLPSEGLS 330

Db 458 LGPLPPGWEERTTDRGVTFNNHNIKKQTQWEDPRMQNV-----AITGPAEP----- 503

Qy 331 EDELPACRYERDLVKLVLRHLSLQOPQAGHCRIVSRREEIFERSYQIMQVRPKD- 389

Db 504 -----YSRIYKRYEFPFRKLKKQTDIPNKEMLKLRANILEDSYRRIMGVKRADF 554

Qy 390 LKRLMYKVFGESEGLDYGGVAREWLYLLCHEMLNPPYGLFQYS-TDNIYMLQINPDSSI- 447

Db 555 LKARLWIEFGESEGLDYGGVAREWFLISKEMFNPPYGLFYSATDNNTLQINPNSGLC 614

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 614

Qy 448 NPDLHLSYFHFVGRIMGLAVFHGHYINGFTVFPYKQLLQKPLQSLDESVDPELHKSIVW 507

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 507

615 NEDHLSYFKFGRVAGMAYVHGKLLDGGFFIRPFYQOMLQKLIITHDMESVDSEYYSRLW 674
508 ILENDITPVLDTFCVEHNAFGRILQHELKPNRNPVPTTEENKEYVRLVYNWRFMEGIE 567
675 ILENDPTE-LDLRFIIDEELFQTHQHELKTPGGSEVVYNNKXKEYVLYVQWRFVNRQ 733
568 APLALQKPNELIQLHLLKPPDQKLELITGLDKIDLNDKSNTRLKHCVA-DSNIVR 626
734 KQWAAPEKGEFFELIPQDLIKIFDENELLMCGLDGVDVNDWRREHKYKNGYSUNHGVH 793
627 WFOVAVETDEERRARLLQFVTSRVRPQLQGFALQSGTGAAGPRLEFTIHLIDANTDNL 686
794 WFKAVLWMDSEKIRLLQFVTSRVRPQMGPAELYGSN---GQSFVTEQW-GTPDKLP 849
687 KAHTCNRRIDIPYSEYKLYKLTAVETCGF 720
850 RAHTCFNRLLDPYSEFDELWDLQWAIENTQGF 883

SULT 8
3196
DD-4 ORF - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
Accession: I83196
Kumar, S.; Tomooka, Y.; Noda, M.
Ochem. Biophys. Res. Commun. 185, 1155-1161, 1992
Title: Identification of a set of genes with developmentally down-regulated expression
Reference number: I80167; PMID:92328780; PMID:1378265
Accession: I83196
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
Residues: 1-708 <RES>
Cross-references: GB:D10714; NID:G220508; PID:G220509
Genetics:
Gene: NEBD-4
Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
40-77/Domain: WW repeat homology <WW1>
196-233/Domain: WW repeat homology <WW2>
251-288/Domain: WW repeat homology <WW3>
347-682/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 27.1%; Score 1051.5; DB 2; Length 708;
Best Local Similarity 36.1%; Pred. No. 2.1e-68;
Matches 244; Conservative 102; Mismatches 194; Indels 135; Gaps 19;
50 LDPKWNQHYDLVYGVKTDITISIVNNHKKIKKQAGFLGCVRLLSNAISRLKXDTGYRLD 109
42 LPPGWEERQDV-LGRYTVV-----NHES-----RRTQWKR-- 70
110 LCKLNP-----DTDAVGGQIVVSLQTRDRIGTGGSVDCRGLLEN-----EGTV 154
71 -----PSPDDDLTDEDNDMDQLQAFRTTRQISDVDPGNRSPENWEIVREDENTE 125
155 YEDSGPRGLSCFME-----BPA---PYTSTGAAAGG--NCRFVES 192
126 YSGAVQSPGSHGIDVQTHLAEBENTELAVCGNPATSPQVTSNHSRSGSLQTCIFEEQ 185
193 PSQORLOAQRNLPDRVGLQTPONRPHGQSPFELPEGVEQRTTVQGVFLHTQTGVS 252
186 PTLFVLLPT-----SSGLPPGWEKQDDRGSRVYVDHNSKTT 222
253 TWHDP-----RIPDL-NSVNCDELGLPPGWEVRSTVSGRIYFVDHNNRTTQFTDP 303
223 TWSKETMODDPRKIPAHLRKTDNSDLGLPFGWEERTHTDGRVFFINHNKKTOWEDP 282
304 RLHHTNHQCLKPSQPLPLPSEGSLEDEBELPAQRVERDLVKLVLRHLSLQOPQAG 363
283 RLQNV-----AITGPAV-----YSRDYKRYKFFFRKLKQTDIPN 319
364 HCRIVSREEIFEEYSRQIMKWRPKD-LKKRLMYKFRGESGLDYGGVAREWLILLCHEML 422
320 KPEMLKRANILLESYRRIMGVKRADLLKARLWIEFPGKGLDYGGVAREWFFLLSKEMF 379

423 NPYGLTQYSTDNIMLOINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHIYNGGFTVPY 481
380 NPYGLTQYSTDNIMLOINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHIYNGGFTVPY 439
482 KQLLGKTIQSLDLESVDPELHKLWLILENDITPVLDTFCVEHNAFGRILQHELKPNR 541
440 KMWLQKLTITLHDMESVDSEYYSRLWILENDPTE-LDLRFIIDEELFQTHQHELKTTGS 498
542 NPYTEENKKEYVLYVYNWRFMRGIEAQLALQKGNELIPOHLLKPPDQKLELITGL 601
499 EIVTNKXKEYVLYVQWRFVNRQKMAAFKEGFEFELIPQDLIKIFDENELLMCGSL 558
602 DKIDLNDKSNTRLKHCVA-DSNIVRWFQAVETPDBERRARLLQVTVGSTRVPLQGFKA 660
559 GDVDVNDWRREHKYKNGYSUNHGVHFWKAVWMDSEKIRLLQFVTSRVRPQMGPAE 618
661 LOGSTGAAGPRLEFTI 675
619 LYGSN---GQSFVTE 630

RESULT 9
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C;Accession: T37900
R;Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21752
A;Accession: T37900
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-671 <RES>
A;Cross-references: EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:SPAC1805.15C
A;Experimental source: strain 972h-; cosmid c1805
C;Genetics:
A;Gene: SPDB:SPAC1805.15C
A;Map position: 1
A;Intons: 60/2; 105/1; 639/2
C;Superfamily: WW repeat homology
F;242-279/Domain: WW repeat homology <WWR>

Query Match 26.4%; Score 1026.5; DB 2; Length 671;
Best Local Similarity 32.9%; Pred. No. 1.3e-66;
Matches 248; Conservative 128; Mismatches 253; Indels 125; Gaps 18;
6 KIRLTVLCAKNLAKDFFRLPDPFAKIVVDGSGQCHSTDTVKNITLDPKWNQHYDLVYVKT 65
7 EVQLTILHVEGLWKNGLRLSLKPYLLISVD-DDQFIKTNVASGTLRLSWGFTQKLTVP 65
66 DSTTISVNNHKKIKKQAGFLGCVRLLSNA---ISRLKDTGYORLDLCKLNPSDTPDAVR 122
66 SILLGLQFDEKQ-KNETSDGFVGLGAAVNVNFPNFKDDYKTRIL-----RSPSGSYR 120
123 GQIVVSLQTRDRIGTGGSVDCRGLLENEGTVYEDSGPRPLSCFMEEPAP-----YTD 176
121 GSVVCLFKR-----SKFLPEELPADKSKQICTD 147
177 STGAAG-----GGNCRFVESP--SQDQRLQAOQLRNPDRVGRSLQTPQNRPHGHQS 225
148 IIDPAGCGAWETRIDFEGHYVYLYKSPQSVISAISHEKLEN-----LTK----- 192
226 PELPEGVQORTTVQGVFLHTQTGVSTW-----HDPRIPLDINSVNCDEL 271
193 -QLKEVFSQ-----FLFNNQSKSLKINLEYKVIKHLLEHYPLALSVRQOVAVEK 241
272 GPLPPGWEVRSTVSGRIYFVDHNNRTTQFTDRLHHIMNHQCLKPSQPLPLPSEGSLE 331
242 GPLPAGWEMELSEDYHYVFDHSTKTITWSDPRDNNV-----ASDSVS 284
332 DEELPAQRVERDLVKLVLRHLSLQOPQAGHCRIVSREEIFEEYSRQIMKWRPKD 389

sen, N.F.; Hughes, B.; Huizar, L.
ture 408, 816-820, 2000
Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
zo, M.; Rooney, T.; Rowley, D.; Sakano, H.
Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
r, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
Reference number: A86141; MUID:21016719; PMID:11130712
Accession: H95599
Status: preliminary
Molecule type: DNA
Residues: 1-4056 <STO>
Cross-references: GB:AE005173; NID:g8778329; PIDN:AAF79338.1; GSPDB:GN00141
Gene: F14J16.10
Map position: 1
Query Match 20.2%; Score 784; DB 2; Length 4056;
Best Local Similarity 39.4%; Pred. No. 9.5e-48;
Matches 172; Conservative 65; Mismatches 138; Indels 62; Gaps 7;
343 DLVOKLVLRHLSLQOPO--AGHCRIEVSREIEPESYRQIMKMRPKDLKRLMVKFRG 400
3624 DFNKKAYFRSRIRHQDHQHSGLPLAISVRRAYVLEDSYNQLMRSPQDLKGRNLNVQFOG 3683
401 EEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTDNIMQLNPSSINPDHLSYHFVGR 460
3684 EEGIDAGGLTREWYQLLSRVIFDKGALLFT-TVGNDATFQPNPNSVYQTEHLSYFKEVGR 3742
461 IMGLAVPHGHVINGGTFVPFYKQLGPKIQLSDLESVDPELHKSIVLWLENDITPVLDDHT 520
3743 MVAKALFDGQLDVFYFRSYKHILGVKTYHDIKAVDPDYKYLKWLLENDVSDILDLT 3802
521 F-----CVE---HNAFGRILQHELPKNGRN 542
3803 FMSDADBEKHILYEKTEVTYDELKPGGRNIRVTEETKHEYVDLVADHILTSAIRPQINAF 945
543 VPVTEENKKEYVLYNWRPMRGIEAQFALOKNGNELIPQHLKLPFOCKELELIIGLD 602
3863 IRVTEETKHEYVDLVAGHILTNAIRPQINAFLEGNELIPRELVSIFNDELLELLISGLP 3922
603 KID-----LNDKSNRLKHCVADSNIRVFWQAVETDEERRARLLQOV 647
3923 EIDCKLDSIOYLCAVPLDLDLANKANTEYSYTAGSPVIHFWVWKAFSKEDMARFLQV 3982
648 TGSTRVFLQGFALQSGTGAAGPRFTIHLIDANTNLPAKTCNFRIDIPPYSEYKLY 707
3983 TGTSKVPLEGFKALQ---GISGQRLQHKAYGAPERLPSAHTCFNQLDLPYQSKQLQ 4039
708 EKLITAVEETC---GFA 721
4040 ERLLLAHEASEGFGFA 4056
RESULT 13
31491
ubiquitin-protein ligase homolog F1707.15 - Arabidopsis thaliana
Species: Arabidopsis thaliana (mouse-ear cress)
Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
Accession: T01491
Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li,
z, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N
bmitted to the EMBL Data Library, June 1998
Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.
Reference number: Z14334
Accession: T01491
Status: translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-1126 <VYS>
Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176690; GSPDB:GN00059; ATSP:F1707
Experimental source: cultivar Columbia
Genetics:

A:Gene: ATSP:F1707.15
A:Map position: 1
A:Introns: 118/3; 224/3; 292/3; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2
C:Superfamily: ubiquitin-protein ligase homology
F:756-1120/Domain: ubiquitin-protein ligase homology <UBI>
Query Match 20.1%; Score 780.5; DB 2; Length 1126;
Best Local Similarity 40.1%; Pred. No. 2.4e-48;
Matches 170; Conservative 67; Mismatches 138; Indels 49; Gaps 6;
343 DLVOKLVLRHLSLQOPO--AGHCRIEVSREIEPESYRQIMKMRPKDLKRLMVKFRG 400
707 DFNKKAYFRSRIRHQDHQHSGLPLAISVRRAYVLEDSYNQLMRSPQDLKGRNLNVQFOG 766
401 EEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTDNIMQLNPSSINPDHLSYHFVGR 460
767 EEGIDAGGLTREWYQLLSRVIFDKGALLFT-TVGNDATFQPNPNSVYQTEHLSYFKEVGR 825
461 IMGLAVPHGHVINGGTFVPFYKQLGPKIQLSDLESVDPELHKSIVLWLENDITPVLDDHT 520
826 MVAKALFDGQLDVFYFRSYKHILGVKTYHDIKAVDPDYKYLKWLLENDVSDILDLT 885
521 FCVEHNAFGRIL-----QHELPKNGRNVPVTEENKKEYVRLYVNWRFMRGIEAQFLAL 573
886 FMSDADBEKHILYEKTEVTYDELKPGGRNIRVTEETKHEYVDLVADHILTSAIRPQINAF 945
574 QXGFNELIPQHLKLPFOCKELELIIGLDKID----- 605
946 LEGLNELIPRELVSIFNDELLELLISGLPEIDCKLAPYFIFLHFAYSFKVITILLSV 1005
606 -----LNDKSNRLKHCVADSNIRVFWQAVETDEERRARLLQOVFGTSTRVFLQGFKA 660
1006 FFCFLVDLDLANKANTEYSYTAGSPVIHFWVWKAFSKEDMARFLQVFGTSTRVFLQGFKA 1065
661 LOGSTGAAGPRFTIHLIDANTNLPAKTCNFRIDIPPYSEYKLYEKLITAVEET--- 717
1066 LQ---GISGQRLQHKAYGAPERLPSAHTCFNQLDLPYQSKQVQERQLLLAHEANEG 1122
718 CGFA 721
1123 GFGA 1126
RESULT 14
S69625
hypotheical protein YDR457w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C:Accession: S69625
R; Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A:Reference number: S69554
A:Accession: S69625
A:Molecule type: DNA
A:Residues: 1-3268 <DIE>
A:Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64910.1; PID:g927738; MIPS:YDR45
C:Genetics:
A:Gene: SGD:TOM1
A:Cross-references: SGD:S0002865; MIPS:YDR457w
A:Map position: 4R
Query Match 19.8%; Score 769; DB 2; Length 3268;
Best Local Similarity 43.3%; Pred. No. 8.4e-47;
Matches 152; Conservative 70; Mismatches 125; Indels 4; Gaps 2;
367 IEVSREIEPESYRQIMKMRPKDLK-KELMVKFRGEGLDYGGVAREWLYLLCHEMLNPY 425
2913 ITVRRQVFLDSYRALFFKTNDEIKNSKLEITFKGESGVADAGVTREWYQLSRQEMND 2972
426 YGLFOYSTDNIMQLNPSSINPDHLSYHFVGRIMGLAVPHGHVINGGTFVYKQLL 485
2973 YALFLVPSDKTTFHPNRTSGINPEHLSFFKPIGMIIGKAIKRDQCFLDCHFSREVYKNI 3032

486 GKPIQLSDLESVDPPELHKSILVILENDITPVLDTHTPCVEHNAFGRILQHELKPNGRNVFV 545
 3033 GRPVSUKMESLDPDYKSLVILENDITDIIEETFSVETDDYGEHKVINLEGKDIIV 3092
 546 TENKKEVRLYNWFMGIEAQFLALOKGNELIPQHLKPPQKLELIIGGLDKID 605
 3093 TEANKQDYKVVVEYKQTSQMDNFLVGFYALISKDLITIFDEQSELLISGLPID 3152
 606 LNDWKSNTLKHCVADSNIVRWFWQAVETFDERRARLLOFTGSTRVPLQGFKALQGST 665
 3153 VDWKANTTVYNTATCKEYSVFWRAVRFDAERAKLLOFTGTSKVPPLNGFKELSGVN 3212
 666 GAAGPRFTTHLIDANTDNLPKAHTCFNRIDIPPYESYEKLYEKLITAVEE 716
 3213 GVC---KFSIHRDFGSSERLPSSHTCFNQLNLPYESTELRGSLLAINE 3260
 ISULT 15
 0274
 pothetical protein F45H7.6 - Caenorhabditis elegans
 Species: Caenorhabditis elegans
 Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
 Accession: T20274; T22257
 Percy, C.
 Submitted to the EMBL Data Library, November 1994
 Reference number: Z19245
 Accession: T20274
 Status: preliminary; translated from GB/EMBL/DDBJ
 Molecule type: DNA
 Residues: 1-889 <W1>
 Cross-references: EMBL:Z46793; PIDN:CAA86773.1; GSPDB:GN00021; CESP:F45H7.6
 Experimental source: clone C56G7
 Percy, C.
 Submitted to the EMBL Data Library, June 1994
 Reference number: Z19538
 Accession: T22257
 Status: preliminary; translated from GB/EMBL/DDBJ
 Molecule type: DNA
 Residues: 1-889 <W12>
 Cross-references: EMBL:Z34800; PIDN:CAA84325.1; GSPDB:GN00021; CESP:F45H7.6
 Experimental source: clone F45H7
 Genetics:
 Gene: CESP:F45H7.6
 Map position: 3
 Introns: 30/3; 54/1; 118/1; 185/2; 213/1; 260/3; 288/2; 353/2; 454/1; 559/2; 642/2; 74
 Superfamily: WW repeat homology
 229-266/Domain: WW repeat homology <WW1>
 372-408/Domain: WW repeat homology <WW2>
 Query Match 17.2%; Score 669; DB 2; Length 889;
 Best Local Similarity 28.7%; Pred. No. 2.2e-40;
 Matches 168; Conservative 104; Mismatches 203; Indels 110; Gaps 15;
 y 232 YQRTTVGQVYFLHTQGVSTWHDPRIPDLN-----SVNCELOPLPP 276
 b 322 YENQDAMQ-----IYNEGSVVRHAIHQDLDPSPKFENQPLFVFNLFADITQPLPS 376
 y 277 GWEVSTVSGRIYFVDHNNRTQTDPRLHNNHQQCKEPPQPLPSEGS----- 329
 b 377 GWEC-ITWNRVFLNHNKETSFYDPRAPF-----ETKTSRGRSVPSRSSTAHKGI 430
 y 330 -----LEDELP--AQYERDLV----- 345
 b 431 DHALISKCEDLRKIAQDNFPQIAERISKMLIERFGGLAVASLANDLITLALSMLDSN 490
 y 346 -----QKLVLRHELSQLQPGAGHCRV--EYSREEIPESYRQIMKWRPKDLAK-RL 394
 b 491 TEKLAGEGDNITKFTYEDMKKEKLGKPSRLCWKVRDRLLDDAPRIILNVDPFVLKKSRL 550
 y 395 MYKFRGEGLDYGGVAREWLYLLCHEMLNPNYIGLFQYSTONIYMLQINPDSS-INPDHLSY 454
 b 551 HIRFEGELALDYGGLSREFFILLSRFLPKNQGYFEVE-GNDYHLQLRPRGCEKEKWK 609

QY 455 FHEVGRIMGLAVFHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPPELHKSILVILENDIT 514
 Db 610 LILGCRVLALAVIHRCYIDVFTTNVFKSLQKRPVTLMDPKESDABFYKSNMILLENDVV 669
 QY 515 FVLDHTFCVEHNAFGRI-----LQHELKPNGRNVFVTEENKKEVRLYNV 559
 Db 670 D-LEMSFVYSSVMWNGKVSDDLTLSSIVSVLALQAOELVPGESQKVTANKAEFIDLNCQ 728
 QY 560 WRFMRGIEAQFLALOKGNELIPQHLKPPQKLELIIGGLDKIDLNDWKSNTLKHCV 619
 Db 729 KKAIRGVEKPELEILITSFNQILNDNLLNSLESSDLKRILSGSLELDLNDWRTNTIYKGGY 788
 QY 620 ADSNI-VRWFWQAVETFDERRARLLOFTGSTRVPLQGFKALQGSTGAACPRFTTHLI 678
 Db 789 SDCHIVVEFWFVEIETMTNQERFDLLLEFVTGSSVFFEGFSALRGNEIS---KFCIEKW 845
 QY 679 DANTONLPKAHTCFNRIDIPPYESYEKLYEKLITAVEETCGFAVE 723
 Db 846 -GDATSFPAHTCFNRLQLPSYNTKQKLSKLOQAIYNGMSYSIE 889

Search completed: February 20, 2004, 15:26:53
 Job time : 20.2026 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

n on: February 20, 2004, 15:25:26 : Search time 10.8131 Seconds
(without alignments)
3144.377 Million cell updates/sec

tle: US-10-009-945-2

fect score: 3884

quence: 1 GGSSIKIRLTVCANLAKK.....EKLYEKLITAVETCGFAVE 723

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 127863 seqs, 47026705 residues

tal number of hits satisfying chosen parameters: 127863

nimum DB seq length: 0

ximum DB seq length: 200000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	3855	99.3	757	1	SUFI1_HUMAN
2	3548	91.3	731	1	SUFI1_XENLA
3	3249.5	83.7	619	1	SUFI1_MOUSE
4	3001.5	77.3	748	1	SUFI2_HUMAN
5	1492.5	38.4	787	1	PUB1_SCHPO
6	1377	35.5	809	1	RSP5_YEAST
7	1259	32.4	927	1	NED4_HUMAN
8	1225	31.5	957	1	NED4_MOUSE
9	627.5	16.2	310	1	URB1_RAT
10	534	13.7	875	1	UE3A_HUMAN
11	483	12.4	885	1	UE3A_MOUSE
12	481.5	12.4	882	1	HUL4_YEAST
13	449.5	11.6	1050	1	HER3_HUMAN
14	358	9.2	1992	1	TRIB_HUMAN
15	354	9.1	910	1	HUL5_YEAST
16	278.5	7.2	2799	1	EDD_HUMAN
17	277	7.1	920	1	EDD_RAT
18	262.5	6.8	1483	1	UFDA_YEAST
19	257	6.6	1647	1	YDEI_SCHPO
20	238.5	6.1	472	1	YAPI_MOUSE
21	238.5	6.1	1620	1	HEDI_HUMAN
22	222	5.7	2895	1	HYD_DROME
23	162	4.2	448	1	YAPI_CHICK
24	161.5	4.2	454	1	YAPI_HUMAN
25	128	3.3	1813	1	UN13_CAEEL
26	119	3.1	575	1	BAG3_HUMAN
27	116.5	3.0	577	1	BAG3_MOUSE
28	114.5	2.9	826	1	RSBG_HUMAN
29	114	2.9	593	1	CNE5_HUMAN
30	113	2.9	1021	1	YPT7_CAEEL
31	111.5	2.9	672	1	KPCA_HUMAN
32	111.5	2.9	672	1	KPCA_RABIT
33	111.5	2.9	672	1	KPCA_RAT

34	111.5	2.9	5147	1	PCLO_HUMAN
35	110.5	2.8	597	1	IXR1_YEAST
36	110	2.8	658	1	KPC1_LYPTI
37	109.5	2.8	537	1	CNE3_HUMAN
38	109	2.8	694	1	RP3A_HUMAN
39	109	2.8	704	1	RP3A_BOVIN
40	109	2.8	915	1	PDB2_ARATH
41	108	2.8	506	1	PYCA_ARCFU
42	108	2.8	665	1	RNR2_LACLA
43	108	2.8	1013	1	PTPX_MACNE
44	108	2.8	1178	1	PYC_MOUSE
45	108	2.8	5120	1	PCLO_CHICK

ALIGNMENTS

RESULT 1
SUFI1_HUMAN STANDARD; PRT; 757 AA.
AC Q9HCE7; O75853; Q9UUT8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
DE (hSMURF1).
GN SMURF1 OR KIAA1625.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).
RX MEDLINE=99385348; PubMed=10458166;
RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RT pattern formation.";
RL Nature 400:687-693(1999).
RN [2]
RP SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
RA Stoneking T., Bauer C., O'Neal D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RT Tissue=Brain;
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroseawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res 7:273-281(2000).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q9HCE7-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q9HCE7-2; Sequence=VSP_006812;
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).

```
-----
EMBL; AF199364; AAF08298.2; -
EMBL; AC004893; AAC62434.1; -
EMBL; AB046845; BAB13451.1; ALT_INIT.
HSSP; Q13526; 1PIN.
MIM; 605568; -
GO; GO:0005622; C:intracellular; TAS.
GO; GO:0002211; P:protein degradation tagging activity; IDA.
GO; GO:0004842; P:ubiquitin-protein ligase activity; IDA.
GO; GO:0000424; P:cell differentiation; TAS.
GO; GO:00030154; P:ectoderm development; IDA.
GO; GO:0007398; P:ectoderm development; TAS.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
GO; GO:0004644; P:protein modification; TAS.
GO; GO:0006512; P:ubiquitin cycle; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT_domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00459; C2_DOMAIN_1; 1.
PROSITE; PS00004; C2_DOMAIN_2; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 1.
PROSITE; PS0020; WW_DOMAIN_2; 2.
Ubl conjugation pathway; Ligase; Repeat; Alternative splicing.
DOMAIN 1 99 C2 DOMAIN.
DOMAIN 234 267 WW 1.
DOMAIN 306 339 WW 2.
DOMAIN 420 757 HECT.
BINDING 725 725 UBIQUITIN.
VARSPPLIC 269 294 Missing (in isoform Short).
MUTAGEN 725 725 C->A; LOSS OF UBIQUITINATION CAPACITY.
SEQUENCE 757 AA; 86113 MW; 89A171CF647B40E9 CRC64;
Query Match 99.3%; Score 3855; DB 1; Length 757;
Best Local Similarity 96.5%; Pred. No. 8.8e-272;
Matches 722; Conservative 0; Mismatches 0; Indels 26; Gaps 1;
2 GSSIKRLTVLCANLAKDFRLPPAKIVVDGSGGCHSDTDTVNTLDPKWNQHYDLY 61
10 GSSIKRLTVLCANLAKDFRLPPAKIVVDGSGGCHSDTDTVNTLDPKWNQHYDLY 69
62 VGKTDSTISVWNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQLDLCKLNPSTDV 121
70 VGKTDSTISVWNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQLDLCKLNPSTDV 129
122 RQIVVSLQTRDIBIGTGGVWDCRGLLENGTVYDSGGRPLSCFMEEPAPYTDSTGA 181
130 RQIVVSLQTRDIBIGTGGVWDCRGLLENGTVYDSGGRPLSCFMEEPAPYTDSTGA 189
182 AGGNCRCFVESPQDORLQAQLRNPDRVRSGLTQPNRPHGHSQSELPEGYEQRTTVQ 241
190 AGGNCRCFVESPQDORLQAQLRNPDRVRSGLTQPNRPHGHSQSELPEGYEQRTTV 249
242 VYFLHTGTGTVTHDPRI-----PDLNSVNCDELGLP 275
250 VYFLHTGTGTVTHDPRI-----PDLNSVNCDELGLP 309
276 PGWEVRVTSGRIVYFVDHNNRTQFTDPRLLHMHQCOLKEPSPQLPSPSGSLEDEL 335
310 PGWEVRVTSGRIVYFVDHNNRTQFTDPRLLHMHQCOLKEPSPQLPSPSGSLEDEL 369
336 PAQRYERDLVQKLVRLHELSLQOPAGHCRIVRSREIFEESYQIMKRPKDLKGLM 395
370 PAQRYERDLVQKLVRLHELSLQOPAGHCRIVRSREIFEESYQIMKRPKDLKGLM 429
396 VKRGEGLDYGGVAREWLYLLCHEMLNPFYGLFOYSTDNLYMLQINPDSSINPDHLSYF 455
-----
```

```
DB 430 VKRGEGLDYGGVAREWLYLLCHEMLNPFYGLFOYSTDNLYMLQINPDSSINPDHLSYF 489
QY 456 HFVGRIMGLAVFHGHINGGFTVPFYKQLLKPQLQSLDLESVDPELHKSILVWLENDITP 515
DB 490 HFVGRIMGLAVFHGHINGGFTVPFYKQLLKPQLQSLDLESVDPELHKSILVWLENDITP 549
QY 516 VLDHTFCVEHNAFGRIILQHELKPNGRNVPVTENKKEYVRLYVNRFRMEGIEAQLALQK 575
DB 550 VLDHTFCVEHNAFGRIILQHELKPNGRNVPVTENKKEYVRLYVNRFRMEGIEAQLALQK 609
QY 576 GFNELIPOHLLKFPDQKELELIIGGLDKDLNDWKSNTLKHCVADSNIVRWFQAVETP 635
DB 610 GFNELIPOHLLKFPDQKELELIIGGLDKDLNDWKSNTLKHCVADSNIVRWFQAVETP 669
QY 636 DEERRARLLQFVTGSTRVPLQGFKAQGGSTGAAGPRLFTIHLIDANTDNLKPAHTCFNRI 695
DB 670 DEERRARLLQFVTGSTRVPLQGFKAQGGSTGAAGPRLFTIHLIDANTDNLKPAHTCFNRI 729
QY 696 DIPPYSEYKLYEKLTAVEETCGFAVE 723
DB 730 DIPPYSEYKLYEKLTAVEETCGFAVE 757

RESULT 2
SUFL_XENLA
ID SUFL_XENLA STANDARD; PRT; 731 AA.
AC O9PUN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Smad ubiquitin ligase regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase) (xSMURF1).
SMURF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
TISUE=Blasutula;
MEDLINE=99385348; PubMed=10458166;
RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
pattern formation."
RL Nature 400:687-693(1999).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2
(ACTIVIN/TGF-BETA) PATHWAY.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING
TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO
GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC
ECTODERM AND INVOLUTING MESODERM, AND EXPRESSION GRADUALLY
LOCALIZES TO THE NERVOUS SYSTEM. AT EARLY TADPOLE STAGES EXPRESSED
IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF169310; RAD52564.1; -
GO; GO:0005622; C:intracellular; TAS.
DR
```

```
GO; GO:0000211; F:protein degradation tagging activity; IDA.
GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO; GO:0030154; P:cell differentiation; IDA.
GO; GO:0007398; P:ectoderm development; TAS.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
GO; GO:0006464; P:protein modification; TAS.
GO; GO:0006512; P:ubiquitin cycle; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Reps_WW.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00239; C2; 1.
SMART; SM00119; HECT; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00499; C2_DOMAIN_1; 1.
PROSITE; PS00004; C2_DOMAIN_2; 1.
PROSITE; PS00237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 1.
PROSITE; PS00020; WW_DOMAIN_2; 2.
UBI conjugation pathway; Ligase; Repeat.
DOMAIN 1
99 C2 DOMAIN.
DOMAIN 233
266 WW 1.
DOMAIN 279
312 WW 2.
DOMAIN 394
731 HECT.
SEQUENCE 731 AA; 83259 MW; 3CE88E512A42CE2C CRC64;

Query March 91.38; Score 3548; DB 1; Length 731;
Best Local Similarity 91.08; Pred. No. 1.5e-249;
Matches 659; Conservative 34; Mismatches 29; Indels 2; Gaps 2;

1 GGSIKIRLTVLCAKLNKADFFRLPDPFAKIVVDGSGGCHSTDTVKNTLDPKMQHYDL 60
9 GGSIKIRLTVLCAKLNKADFFRLPDPFAKIVVDGSGGCHSTDTVKNTLDPKMQHYDL 68
61 YVGKTDITTSVNHKKIHKQAGFLGCVRLSNLSAISRKTGTGVRDLCKLNSDTDA 120
69 YVGKTDITTSVNHKKIHKQAGFLGCVRLSNLSAISRKTGTGVRDLCKLNSDTDA 128
121 VVGKTDITTSVNHKKIHKQAGFLGCVRLSNLSAISRKTGTGVRDLCKLNSDTDA 180
129 VVGKTDITTSVNHKKIHKQAGFLGCVRLSNLSAISRKTGTGVRDLCKLNSDTDA 187
181 AAGGNCRFVSPQDQRLQARLNPVRSGLQTPNPHGSHQSPQLPEGVQRTTQVQ 240
188 AAGGNCRFVSPQDQRLQARLNPVRSGLQTPNPHGSHQSPQLPEGVQRTTQVQ 247
241 QVYFLHTQTVSTWHDPRIPDLNSVNCDELGLPGLPGWEVRSTVSGRIYFVDHNNRTQF 300
248 QVYFLHTQTVSTWHDPRIPDLNSVNCDELGLPGLPGWEVRSTVSGRIYFVDHNNRTQF 307
301 TDLRLHMHQCOLKPSQPLPSPSGSLD-BELPAQRYERDLVQKLVLRLHLSLQ 359
308 TDLRLHMHQCOLKPSQPLPSPSGSLD-BELPAQRYERDLVQKLVLRLHLSLQ 367
360 POAGHCRVRSREIPESVQIMKMPKDKRLVMKFRGEGGLDYGVAWEWYLLCH 419
368 POAGHCRVRSREIPESVQIMKMPKDKRLVMKFRGEGGLDYGVAWEWYLLCH 427
420 EMLNPFYGLFYSTNTYMLQNPDSINPDHLSYHFVGRIMGLAVFHGYNGFTVP 479
428 EMLNPFYGLFYSTNTYMLQNPDSINPDHLSYHFVGRIMGLAVFHGYNGFTVP 487
480 FYKQLGKPIQLSDLESVDPELHSLVWILENDITVLDHFTCVENAGRILOHLEKPN 539
488 FYKQLGKPIQLSDLESVDPELHSLVWILENDITVLDHFTCVENAGRILOHLEKPN 547
540 GRNVPVTENKKEYVRLVYVNRFRGTEAQLQKGFNELIPQHLKXPFQKLELIIG 599
548 GKNLQVTEENKKEYVRLVYVNRFRGTEAQLQKGFNELIPQHLKXPFQKLELIIG 607
600 GLDKIDLNDKSNTRLKHCVDASINIVFWQAVETDEERRARLLQFVTGSTRVPLQCFK 659

Db 608 GLDKIDLNDKSNTRLKHCVDASINIVFWQAVETDEERRARLLQFVTGSTRVPLQCFK 667
Qy 660 ALQGSTGAAGPRLFTIHLIDANTDNLKPAHCFNRPIDIPYSEYKLYKLLTAVERTCG 719
Db 668 ALQGSTGAAGPRLFTIHLIDANTDNLKPAHCFNRPIDIPYSEYKLYKLLTAVERTCG 727
Qy 720 FAVE 723
Db 728 FAVE 731

RESULT 3
ID SUFL MOUSE STANDARD; PRT; 619 AA.
AC QSCUN6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
DE (Fragment).
GN SMURF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
STRAN=C57BL/6J; TISSUE=Testis;
MEDLINE=21085660; PubMed=11217851;
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Gustinstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Brownstein S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Maehama J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.;
"Functional annotation of a full-length mouse cDNA collection.";
Nature 409:685-690(2001).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BNP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION (BY
CC SIMILARITY).
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AK015264; BAB29770.1; -
CC HSP; Q13526; IPIN.
CC MGD; MGI:1923038; 4930431E10Rik.
CC GO; GO:0005622; C:intracellular; ISS.
CC GO; GO:0000211; F:protein degradation tagging activity; ISS.
CC
```


	InterPro; IPR000008; C2	
	InterPro; IPR000589; HECT domain.	
	InterPro; IPR001202; WW_Rsp5_WWP.	
	Pfam; PF00168; C2; 1.	
	Pfam; PF00632; HECT; 1.	
	Pfam; PF00397; WW; 3.	
	SMART; SM00239; C2; 1.	
	SMART; SN00119; HECTC; 1.	
	SMART; SM00456; WW; 3.	
	PROSITE; PS00499; C2_DOMAIN_1; 1.	
	PROSITE; PS00004; C2_DOMAIN_2; 1.	
	PROSITE; PS00237; HECT; 1.	
	PROSITE; PS01159; WW_DOMAIN_1; 1.	
	PROSITE; PS00020; WW_DOMAIN_2; 3.	
	Ubl conjugation pathway; LiGase; Repeat; Nuclear protein.	
	C2 DOMAIN.	
	DOMAIN 1 98	
	DOMAIN 157 190	
	DOMAIN 251 284	
	DOMAIN 297 330	
	DOMAIN 414 748	
	MUTAGEN 251 284	
	MUTAGEN 297 330	
	MUTAGEN 716 716	
	MUTAGEN 716 716	
	MUTAGEN 716 716	
	CONFLICT 6 6	
	SEQUENCE 748 AA; 86195 MW; 3042B43A3755762 CRC64;	
	Query Match	
	Best Local Similarity 77.3%; Score 3001; DB 1; Length 748;	
	Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;	
/	5 IKRLTVLCARNLAKYDFRLPDPRAKIVDGGSGCHSTDTVKNTLDPKNQHYDLVYCK 64	
/	12 VKRLTVLCARNLAKYDFRLPDPFAKVVDGGSGCHSTDTVKNTLDPKNQHYDYLYIKR 71	
/	65 TDSITISWNHKKIHKQGAGFLGCVRLLSNALSKDVTGYQLDLCKLPSPSTDVAQQ 124	
/	72 SDSVTISWNHKKIHKQGAGFLGCVRLLSNAINELKDTGYQLDLCKLGENDNDAVQQ 131	
/	125 IVWSLTDRDIGTGGSVDCRGLENE-----GVVY----- 155	
/	132 IWLSQGRDRIGTGQGVDCSRSLFNDLDPGWERTASGRIOYLNIHTTTQWERPTRP 191	
/	156 --EDSGPGRLSCFMERPAPTYDSTGAAGGCCNCFVESPQDQRLAQRLNPFDVRSGL 213	
/	192 ASEYSSPGRLSCFVDENTISGTNGATCG-----QSSDPLAEARRVRSQRHNRYM 243	
/	214 QTPQNRRPHGSQSPPELPGYQRTTVGGYVFHTQTGSTVHDHPRIFRDINSVNCDELGP 273	
/	243 ----SRTLHLTPDLPGEGYQRTTQGGYVFHTQTGSTVHDHPRVFRDLNSINCEELGP 298	
/	274 LPPGWVEASTVSGRIYFVDHNHNTTQFTDPR----LHHIMNHQCLKEPSQPLPLPSEGS 329	
b	299 LPFGWEIRNTATGRVYFVDHNHNTTQFTDPRLSANHLVLNRQKLKDQQQQVV---S 354	
/	330 L---EDELPAQRVERDLVKLVLRHELSQLPOAGHCRIEVSREEIFESYRQIMKWR 386	
b	355 LCDDTECLTVPYRKDLVKLILSQELSQQOQAGHCRIEVSREEIFESYRQVMKWR 414	
/	387 PKDLAKRLMWAFREEGLDYGVARWLVLLCHEMLNPYGLFOYSTDNITYMLQINPDSS 446	
b	415 PKDLAKRLMIKFREEGLDYGVARWLVLLSHEMLNPYGLFYSDRDITYLQINPDSA 474	
/	447 INPDLSYFHFPVRIMGVLAVFHGHYINGGTFVPFYKQLLGKPIQLSDLESVDLKHSLV 506	
/	475 VNPELSYFHFPVRIMGMAVFNHGHIYIDGGFTLPFFYKQLLGKSIITLDLMELVDPLANSLV 534	

Cerrutti L., Lowe T., McCombie W.R., Paulsen J., Potashkin J., Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 "The genome sequence of *Schizosaccharomyces pombe*."
 Nature 415:871-880(2002).
 -1- FUNCTION: REGULATES UBIQUITINATION OF CDC25.
 -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP + diphosphate + protein N-ubiquityllysine.
 -1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-thiolester formation.
 -1- SIMILARITY: Contains 1 C2 domain.
 -1- SIMILARITY: Contains 3 WW domains.
 -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
 -1- SIMILARITY: STRONG, TO YEAST RSP5.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; Y07592; CA68867.1; -;
 EMBL; U66716; AAB07534.1; -;
 EMBL; Z99161; CAB16207.1; -;
 EMBL; U62795; AAB63350.1; -;
 PIR; S66562; S66562.
 PIR; T37545; T37545.
 HSSP; Q13526; 1PIN.
 GeneDB SPombe; SPAC11G7.02; -;
 InterPro; IPR000008; C2;
 InterPro; IPR000569; HECT_domain.
 InterPro; IPR002349; WW.
 InterPro; IPR001202; WW_Rsp5_WWP.
 Pfam; PF00168; C2; 1.
 Pfam; PF00632; HECT; 1.
 Pfam; PF00397; WW; 3.
 PRINTS; PR00403; WWDOMAIN.
 SMART; SM00239; C2; 1.
 SMART; SM00119; HECTC; 1.
 SMART; SM00456; WW; 3.
 PROSITE; PS00499; C2_DOMAIN_1; 1.
 PROSITE; PS50004; C2_DOMAIN_2; 1.
 PROSITE; PS50237; HECT; 1.
 PROSITE; PS01159; WW_DOMAIN_1; 3.
 PROSITE; PS50020; WW_DOMAIN_2; 3.
 UBL conjugation pathway; Ligase; Repeat.
 DOMAIN 17 32
 C2 DOMAIN.
 DOMAIN 211 236
 WW 1.
 DOMAIN 242 247
 POLY-ALA.
 DOMAIN 294 319
 WW 2.
 DOMAIN 351 376
 WW 3.
 DOMAIN 463 767
 HECT.
 BINDING 735
 UBIQUITIN (BY SIMILARITY).
 Q -> K (IN REF. 1).
 CONFLICT 163 609
 MISSING (IN REF. 1).
 CONFLICT 661 661
 T -> K (IN REF. 1).
 SEQUENCE 767 AA; 87267 MW; F1455A155EB9ACF7 CRC64;
 Query Match 38.4%; Score 1492.5; DB 1; Length 767;
 Best Local Similarity 41.1%; Pred. No. 2.1e-100;
 Matches 329; Conservative 118; Mismatches 234; Indels 119; Gaps 19;
 4 SIKRITLVCAKLNKDFRPPDPFAKTVDPGSGCHSTDTVKNLDPKNQHYDLYVG 63
 7 SRRIRTVIAADGLYKEDVRFPDPFAVTVDGE-QTHTTAKTLNPNYNEFVNV 65
 64 KTSITISVNHKKIHKQAGPLGCVRLLSNAIRLKDVTGQRL--DLCKLNPSDITAV 121
 66 DNSTIAIQVDFQKFK-KKRGQGLGVNLRVGDVLDLAIGDMDLTRLKKS--ENTV 122
 122 RGQIVWSLQ-----TRDRTGTGSGVDCRGLLENGTVYED 157

Db 123 HGKIIINLITTAQSTLQVPSAASGARTQRTSTINDPQSSQSSSVSRNPASSRAGSPFD 182
 QY 158 SGPG-----RPLSCFMBE-----PAPY-----TDSTG-----AAA 182
 Db 183 NAPAASPASSEPRTFSSFDQYGRLPFGWERRTDNLGRTVYVDHNRSTTWIRPNLSVA 242
 QY 183 GGGCRPFVSPS-----QDQRLQARLRNPDVRGSLQTPQNRPHGHSPELPGYEQ 234
 Db 243 GAAAEALHSSASSANVTGQVPSSSNARTEASVLTSTNATTAG-----SGELPFGWEQ 296
 QY 235 RTTVQGVYFLHTQTGVSTVHWDPRIPDLNSVN-----CDELPLPGWVRS 283
 Db 297 RYTPGRPYFVDHNRTRTTTWDPERRQYIRSYGPNNAITQQQPVSQLPLPSGWEMLT 356
 QY 284 VSGRTYFVDHNRTRTTQDRLHHVNHQCLKEPSQPLPSEGSLEDEELPAQRYERD 343
 Db 357 NTRVYFVDHNRKTTTWDPR-----LPS--SL-DQVDP--QYKRD 392
 QY 344 LVQKLKVLRLHLSLQQAGHCRIEYSREIPEESYRQIMQMRPKDLKRLMYKFRGEG 403
 Db 393 FRKLIYFLSQPAL-HPLPGQCHIKVRNHFEDSYAEIMRQSATDLKRLMIKPDGEG 451
 QY 404 LDYGVAREWLYLICHEMLNPPYCLFOYSTDNTYMLQINPDSSINPDHLSYHEFVGRMG 463
 Db 452 LDYGLSREYFFLLSHWFNFYCLFEYSVDNTYQINPHSGINFEHLNYFAFIRGVIG 511
 QY 464 LAVFHGHIINGFTVPFYKQLGKPIQLSDLESVDPELHKSLVWILENDITPYLDHTFCV 523
 Db 512 LAIFRRFVDAFFVYFYKMLQKVLQDMSMDAEYRSVLWILDNDITGVLDLTFVS 571
 QY 524 EHNAFGRILQHELKPNRNPVTENKKEVYRLVYVWRPMRGTEAQFLALQKFNELIQ 583
 Db 572 EDNCFGEVVTIDLPNGRNTEVTEENKREYVDLVTVRIQKRIEQFNAPHEGFSLEIPQ 631
 QY 584 HLLKPFDOKELELIIGGLDKIDLNDKSNTRLKHCVADSNIVRFWQAVETFEERRARL 643
 Db 632 ELINVFDERELELLIGSIEDMDWKDTHDYSYSENDQIKWFELMDENSEKKSRL 691
 QY 644 LQVTGSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTNLPAKHTCFNRIDIPYESY 703
 Db 692 LQFTTGTSTRIPVNGFKDQSGD--GPRKFTIEKA-GEPNKLPKAHTCFNRDLPLPYTSK 747
 QY 704 EKLVEKLLTAVEETCGFAVE 723
 Db 748 KDLDEKLSIAVEETIGFQGE 767
 RESULT 6
 RSP5 YEAST
 ID RSP5 YEAST STANDARD; PRT; 809 AA.
 AC P39940.
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase RSP5 (EC 6.3.2.-)
 GN RSP5 OR NP11 OR MDP1 OR YER125W OR SYGP-ORF41.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RC STRAIN=S288c / AB972;
 RC PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 Hunkeler-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V.";
 RT

Nature 387:78-81 (1997).
[2] IDENTIFICATION.
Winston F.;
Unpublished observations (FEB-1993).
[3] CHARACTERIZATION.
STRAIN=Sigma 1278B;
MEDLINE=96154942; PubMed=8596462;
Hein C., Springael J.-Y., Volland C., Hagenauer-Teaplis R., Andre B.;
"Nfil, an essential yeast gene involved in induced degradation of Gap1
and Fur4 permeases, encodes the Rsp5 ubiquitin-protein ligase.";
Mol. Microbiol. 18:77-87 (1995).
[4] FUNCTION.
MEDLINE=95223981; PubMed=7708685;
Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M.;
"A family of proteins structurally and functionally related to the
E6-AP ubiquitin-protein ligase.";
Proc. Natl. Acad. Sci. U.S.A. 92:2563-2567 (1995).
[5] ERATUM.
MEDLINE=95281634; PubMed=7761480;
Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M.;
Proc. Natl. Acad. Sci. U.S.A. 92:5249-5249 (1995).
[6] CHARACTERIZATION.
STRAIN=FY56;
MEDLINE=99077972; PubMed=9858558;
Wang G., Yang J., Huibregtse J.M.;
"Functional domains of the rsp5 ubiquitin-protein ligase.";
Mol. Cell. Biol. 19:342-352 (1999).
-!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES.
REQUIRED FOR UBIQUITINATION AND THEREFORE DEGRADATION OF SEVERAL
CELL SURFACE PROTEINS LIKE GAP1, FUR4, MAL61 AND STE2. ALSO ACTS
ON RSP1.
-!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
-!- PTM: THE UBIQUITINATION APPEARS TO BE THE RESULT OF AN
INTRAMOLECULAR TRANSFER OF UBIQUITIN.
-!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
thioester formation.
-!- SIMILARITY: Contains 1 C2 domain.
-!- SIMILARITY: Contains 3 WW domains.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-!- SIMILARITY: STRONG, TO S.POMBE PUB1.
-!- SIMILARITY: TO YEAST YKL010C.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: U18916; AAC03223.1; -;
PIR: S43217; S43217.
HSSP: Q13526; IPIF.
SGP: S0000927; RSP5.
GO: GO:0000151; C:ubiquitin ligase complex; IDA.
GO: GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO: GO:0006333; P:chromatin assembly/disassembly; IMP.
GO: GO:0006513; P:protein monoubiquitination; IDA.
GO: GO:0000209; P:protein polyubiquitination; IDA.
InterPro: IPR000008; C2.
InterPro: IPR000569; HECT domain.
InterPro: IPR002349; WW.
InterPro: IPR01202; WW_Rsp5_WWP.
Pfam: PF00168; C2.1.
Pfam: PF00632; HECT; 1.
Pfam: PF00397; WW; 3.
PRINTS; PR00403; WWDOMAIN.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS00200; WW_DOMAIN_2; 3.
KW Ubi conjugation pathway; Ligase; Repeat.
FT DOMAIN 1 88
FT DOMAIN 229 262
FT DOMAIN 315 322
FT DOMAIN 331 364
FT DOMAIN 387 420
FT DOMAIN 705 809
FT BINDING 777 777
FT VARIANT 733 733
FT MUTAGEN 777 777
FT SEQUENCE 809 AA; 91816 MW; 6F1836384479E70F CRC64;
Query Match 35.5%; Score 1377; DB 1; Length 809;
Best Local Similarity 37.0%; Pred. No. 5.5e-92;
Matches 313; Conservative 131; Mismatches 238; Indels 164; Gaps 24;
QY 3 SSIKRLTVLCAKNIKAKDFRLPPFAKIVVDGSGOCHSTDTVKNTLDPKWNOHYDL-Y 61
DB 3 SSIIVKVL--VAESLYKRDVFRSPDPFAVLIDGY-QTKSTSAKKTILNFWNETFKFDD 59
QY 62 VKGTDSITISVNHKKIKKQAGFLGCVRL-LSNAISRLKD-----TGQVR-----LDL 110
DB 60 INENSILITIQVEDQKFF-KKKDQGFGLVNVVAVGVGLHLEDATSSGRPREETIRDL 118
QY 111 CKLNPSDTDAVVGQIVW-----SLQTRDRIGTGGSVVDCRGLL 148
DB 119 KKSND--DGMVSGRLIVLVLSKLPSSPHSQAPSGHTASSNTSTTRTNGHSTSS---T 173
QY 149 ENEGTVYEDSGFRLSCFMEEPAP-----YDSTGAAA-----182
DB 174 RNHSTSHPSRGTAQAVESTLQSGTTATNTATTHRSSTNSTSTSATQVSSPEDQYGRLLP 233
QY 183 -----GGCNCRFVE-----SPSQDQLQAORLNPEDVSGSLQTPNR-----PHG 222
DB 234 GWERPTDNFRTYVVDHNTTRTTWKRPDLQ-TEABRGNQLNANTELERRORHRTLPGG 292
QY 223 HQS-----PELPEGYEQRITTVQGVYFLHT 247
DB 293 SSDNSVTYVGGGSIIPVNGAAAAAFAATGTTSGELPSCWGEQRTPEGRAYFVDH 352
QY 248 QTGVSTWHDPRIPDLNSVN-----CDEGLPLPGWEVSTVSGRIYFVDHNRRT 297
DB 353 NRTTDTVDVPRQVIRTYGPTNTTIOQPVSQLGSLPSPGWEMELTNTARYFYVDHNTKT 412
QY 298 TQPTDRLHIMHQCQLKPSQPLPSEGLSEDELPQORVERDLVQKLKLRHLSL 357
DB 413 TWDDDR-----LPS--SL-DQVNP--QYKDFRKYVYFRSQPAL 448
QY 358 QPQAGHCHIEVSREIEFESYRQIMKMRPKDLKRLMKVPRGEGLDYGVAREWLYLL 417
DB 449 -RLPGQCHIKVRRKNIFFEDAYQEIIMRQTPDLKRLMKIPDGEGLDYGVSREFFLL 507
QY 418 CHEMLNPFYGLFOYSTDNILYMLQINPDSSINPDHLVYFHFVGRIMGLAVFHGHVINGFT 477
DB 508 SHEMFNFFCLFEYSAYDNYTIQINPNSGINPEHLNFKFICRVVGLGVFHRRLDAFFV 567
QY 478 VPFYKQLGKPIQLSDLESVDPELHSLVWILENDITPVLDPHTFCVRENAFGRILQHELK 537
DB 568 GALYKMWLRKKVVLQDMEGVDAEVYNSLNWMLNSIDGLDLTFSADDERFGEVTVDLK 627
QY 538 PNGRNVPTTEENKEYVRLYVNWRFMRGIEAQFLALQKGFNELIPQHLLKFFDQKELELI 597


```

628 PDGRNIEVDGNKKEVELYQWRIIVDRVQEQFKAFMGDFGNELIPDLVTVDFERELELL 587
629
598 IGLDKIDLNDKNSRLKHKCVADSNIVRMFWQAVETFDERRARLLQFVTGSTRVPLQG 657
629
668 IGGIAEIDIEDKKHTDVRGQESDEVIONFWKCVSEWDEQARLLQFTTGTSTRIPVNG 747
629
658 FKALOGSTGAAGPLFTTHLIDANTDNLKHAHTCFNRIDIPPYEYSEKLYELLTAVEET 717
629
748 FKDLQESD---GPRRFTIEKA-GEVQQLPKSHTCFNRVLDLPQYVDYDSMKQKLTAVEET 803
629
718 CGFAVE 723
804 IGFQGE 809

34 HUMAN
NED4 HUMAN STANDARD; PRT; 927 AA.
P46934;
01-NOV-1995 (Rel 32, Created)
01-NOV-1995 (Rel 32, Last sequence update)
15-SEP-2003 (Rel 42, Last annotation update)
NEDD4 protein (EC 6.3.2.-) (Fragment).
NEDD4 OR KIAA0093.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Bone marrow;
MEDLINE=95308325; PubMed=7788527;
Nagase T., Miyajima N., Tanaka A., Suzuki T., Seki N., Sato S.,
Tabata S., Ishikawa K.-I., Kawabayashi Y., Kotani H., Nomura N.;
"Prediction of the coding sequences of unidentified human genes. III.
The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by
analysis of cDNA clones from human cell line KG-1.";
DNA Res. 2:37-43(1995).
-1- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES (BY
SIMILARITY).
-1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
thioester formation.
-1- SIMILARITY: TO NEDD-4 PROTEIN.
-1- SIMILARITY: Contains 1 C2 domain.
-1- SIMILARITY: Contains 4 WW domains.
-1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; D42055; BAA07655.1; -
HSSP; Q13526; IPIN.
Genew; HGNC:7727; NEDD4.
MIM; 602278; -.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00188; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 4.
PRINTS; PR00360; C2DOMAIN.
SMART; SM00239; C2; 1.
SMART; SMC0119; HECTC; 1.
SMART; SMC0456; WW; 4.

```

```

DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS0159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 4.
KW Ub1 conjugation pathway; Ligase; Repeat.
FT NON_TER 1 1
FT DOMAIN 4 8 POLY-SER.
FT DOMAIN 48 134 C2_DOMAIN.
FT DOMAIN 212 215 POLY-GLN.
FT DOMAIN 218 251 WW 1.
FT DOMAIN 375 408 WW 2.
FT DOMAIN 448 481 WW 3.
FT DOMAIN 500 533 WW 4.
FT DOMAIN 821 927 HECT.
FT BINDING 894 894 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 927 AA; 107099 MW; CDSB57CD34114259 CRC64;

Query Match 32.4%; Score 1259; DB 1; Length 927;
Best Local Similarity 33.6%; Pred. No. 2.4e-83;
Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

QY 7 IRLTVLCAKNAKKDFRLDPFAKIV---VDGSGQCHSTDTVKNTLDPKW----- 55
Db 48 VRRVVIAGIGLAKDILGASDPVVRVTLVDPVNGVLTSTVQTKIKSLNPKNNEILLFRV 107
QY 56 ---QH-----YDLVYVKTD--SITISVNV-----HKTIHKQK 83
Db 108 HPQOHRLLFVFDENRLTRDDFLGQVDVPLVPLTENPRLEPYTFKDFVLHPRSHKSRV 167
QY 84 AGELGCVRLLSNAISRLKDT-----GYQRLD-----LCKL----- 113
Db 168 KGYL-----RLKMTYLPKTSGEDDNAEQAELEPGVVLDDQDACHLQQQOE 216
QY 114 -----NPSD--TDVARGQIVVSLQ-----TRDR 134
Db 217 PSLPPGWEERQDILGRTYVNVHESRRQWKRPDQNLTDANGNIQLQQAQRAFTTRQ 276
QY 135 IGTGGSVDCGLLEN-----EGTVYEDSG--PGRPLSCFWEPAPTYD----- 176
Db 277 ISEBESVDNQSSENWEIREDTEATYSSQAFPPSPPSNLNDVPHLABELNARLTIFG 336
QY 177 ---STGAAAGGNCR-----FVESP-----QDRLQAQRL-- 204
Db 337 NSAVSQPASSNHSRRGSLQAYTFEEQPTLPVLLPTSSGLPPGWEKQDERGRSYVDH 396
QY 205 -----RNPVRCGLTPQ-----NRPHGQSPELPEGYEQ 234
Db 397 NSRTTTWTKPTVQATVETSQ--TSQSASGASQSAQSTSDGQQVTPQSEIEQGLPKGWEV 456
QY 235 RTTVQGVYFLHTGTGVSTWHDPR--IPRLNSV---NCDELGLPPGWEVRSTVSGRI 288
Db 457 RHAPNGRPFIDHNTKTWEDPRLKIPALHKGKTSLDTSNDLGLPPGWEERTHTDRI 516
QY 289 YFVDHNNTTFTDPRLHHMHNCQKLEPSQLPLPSEGLSEDELPQAPRYERDLVQKL 348
Db 517 FYINHNKRTQWEDPRLNV-----AITGPAV-----YSRDYKRY 553
QY 349 KYLRHLSLQOPQAGHCRIEVSREEIFEESSVROIIMKVRPKD-LKKRLMVKFRGEEGLDYG 407
Db 554 EPRFKLKKQNDIFNKEMKLRRATVLEDSYRIMGVGRADFLKARLWIEFFDGEKGLDYG 613
QY 408 GVAREWLYLLCHEMLNPPYGLFOYSTDNIMYLNQINPDSSI-NPDHLSYFHFVGRINGLAV 466
Db 614 GVAREWFFLISKEMFNPPYGLFEYSATDNTYTLQINPNSGLCNEDHLSYFKFGRVAGMAV 673
QY 467 FGHVINGGFTVPFKOLLGPKIQLSDLESVDLDELKSLVWILENDITPVLDHTFCVHN 526
Db 674 YHKULDGFFTRFPYKMLHAKPITLHDWESVDSEYNSLRWILENDPTE-LDURFIDSE 732
QY 527 AFRGRLQHELPNGRNVPTBENKKEYVRLYVNVNRMFGIEAQFLAQKGFNELIFQHL 586
Db 733 LFGQTHQHELNKGGSEIWTNKNKKEYVILVIQWRFVNRIRQMAAFKEGFELIPQDLI 792

```

557 KPFOKELELIIGLIDKIDLNDKSNTRLKH-CVADSNIVRWFWQAVETEDERRARLLQ 645
 793 KIFDENELELMCGLDGVNDVNDREHTKYXGYSANFQVQWFWKAVLWMDSEKIRLLQ 852
 646 FVTGSTRVPLGFKALQSGTGAAPRFLTHLIDANTDNLKPKAHTCFNRIIDIPPYVESYEK 705
 853 FVTGSTRVPMNGFAELYGSN---GQSFTEQW-GTPEKLPRAHTCFNRLDIPPYVESFEE 908
 706 LYEKLLTAVETCTGF 720
 903 LWDKLOMAIENTQGF 923

ULT 8

4 MOUSE

STANDARD; PRT; 957 AA.

P46935;
 01-NOV-1995 (Rel. 32, Created)
 01-OCT-1996 (Rel. 34, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 NEDD-4 protein (EC 6.3.2.-) (Fragment).
 NEDD4 OR NEDD-4 OR NEDD4A.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1] _SEQUENCE FROM N.A.
 TISSUE=Embryo.
 MEDLINE=92328780; PubMed=1378265;
 Kumar S.; Tomooka Y.; Noda M.;
 "Identification of a set of genes with developmentally down-regulated
 expression in the mouse brain."
 Biochem. Biophys. Res. Commun. 185:1155-1161(1992).
 [2]

REVISIONS.

Kumar S.;
 Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
 AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
 THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES (BY
 SIMILARITY).
 -!- FUNCTION: INVOLVED IN THE EMBRYONIC DEVELOPMENT AND
 DIFFERENTIATION OF THE CENTRAL NERVOUS SYSTEM.
 -!- TISSUE SPECIFICITY: BRAIN.
 -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 thiolester formation.
 -!- SIMILARITY: Contains 1 C2 domain.
 -!- SIMILARITY: Contains 3 WW domains.
 -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; D85414; BAA12803.1; -
 HSSP; Q13526; 1PIN.
 MSD; MGI:97297; Nedd4.
 GO; GO:0005829; Cytozol; IDA.
 GO; GO:0000151; C:ubiquitin ligase complex; IPI.
 GO; GO:0005515; F:protein binding activity; IPI.
 GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
 InterPro; IPR000008; C2.
 InterPro; IPR000569; HECT domain.
 InterPro; IPR001202; WW_Rsp5_WWP.
 Pfam; PF00168; C2, 1.
 Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 3.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00119; HECT; 1.
 DR SMART; SMO0456; WW; 3.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS00004; C2_DOMAIN_2; 1.
 DR PROSITE; PS0237; HECT; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 3.
 DR PROSITE; PS00020; WW_DOMAIN_2; 3.
 KW Ubl conjugation pathway; Ligase; Repeat.
 FT NON TER 1
 FT DOMAIN 150 236 C2 DOMAIN.
 FT DOMAIN 319 352 WW 1.
 FT DOMAIN 475 508 WW 2.
 FT DOMAIN 530 563 WW 3.
 FT DOMAIN 851 957 HECT.
 FT BINDING 924 924 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 957 AA; 109967 MW; 79D5F1DF0187F28 CRC64;
 Query Match 31.5%; Score 1225; DB 1; Length 957;
 Best Local Similarity 33.6%; Pred. No. 7.5e-81;
 Matches 291; Conservative 123; Mismatches 238; Indels 214; Gaps 23;
 QY 7 IRLTVLCAKVLAKDFFRLDPDPKIV---VDGSGQCHSTDTVKNTLDPKWNQHY----- 58
 DB 150 VRKVIAGIAGLAKDILGASDPYRVTLVDYDPMSGILTSVQTKIKSLNPKWNEILFRV 209
 QY 59 ---DLVYVKTD--SITISVMN-----HKKIKKQG 83
 DB 210 LPQHRILPEVFDENLRTDDFLGQVDVPLPLPTENPRMERPYTFKDFVLHPRSHKSRV 269
 QY 84 AGFLGCVRLLSNAISRLKOT----- 103
 DB 270 KGYL-----RLKMTYLPKNGSEDENADQAELEFGVWVLDQDPAATHLPHPEP 318
 QY 104 ---GY-QRLDLC-----KLNPSDTDAVRGQIVVSLQTRDRI 135
 DB 319 SPLPGWEERQDVLGRYYVNHESRRTQWKRPSPDDLLTDEDNDMLQLOAQAFTTRQI 378
 QY 136 GTGGSVVDCRGLLEN-----EGTVYEDSGPGRPLSCFME-----EP 171
 DB 379 SEDVDGPDNRESPENWEIVREDENTEYSQAVSQSPESHIDVQTHLAEFNLAVCGNP 438
 QY 172 A---PYTDSGTGAAGG--NCRVESPSQDRLQAQRLANPDVRGSLQTPQNRPHGHQSP 226
 DB 439 ATSPVTSSNHSRGGSLQTCIFEQFTLPVLLPT-----SS 475
 QY 227 ELPEGYQRTTVGQVYFLHTQTGVSTWHD-----RIPRDL-NSVNCDELGLPLPG 277
 DB 476 GLPPGWEKQDDRGRSYVDHNSKTTTWSKPTMQDDPSRKIPAHLRGKTDSNDLGLPLPG 535
 QY 278 WEVSTVSGRIYFVDHNRRTQFTDRLHHNNHCQLAKEPSPQLPLPSEGSLEDEELPA 337
 DB 536 WEERTHTDGRVFFINHNKTKTQWEDPRLQNV-----AITGPAVP----- 574
 QY 338 QRYERDLVQKLKVLRLHLSLQPFQAGHCRIEVSREIFEEYRQIMMPEKD-LKKRLMY 396
 DB 575 --YSDYKRYEYFERRKLLKQTDIPNKFEMKLARANILEDYSYRIMGVRAADILKARLWI 632
 QY 397 KFRREGLDYGVAREWMLYLLCHEMLNPYGLFQYSTDNIMQLINPDSSI-NPDHLSYF 455
 DB 633 EFDGEKGLDYGVAWEFFLISKEMFNYPYGLFEYSATDNYTLQINPNSGLCNEDHLSYF 692
 QY 456 HFVGRIMGLAVFHGHYINGGFTVPFYKQLGKFIQLSDLESVDPELHKSILWILENDITP 515
 DB 693 KFIQGVAGMAYVHGKLLDGGFIRPFYKMWLQKLTLDHMEVDSEYSSLRWILENDPTE 752
 QY 516 VLDHTFCVHNNAFGRILOHELKPNGRNVPTVEENKKEYVRLVYNNFMFGIEAQFLALQK 575
 DB 753 -LDLRFIIDEELFGQTHQHELKTGGSEIVVTNKNKKEYLYLVQWFEFVARIQKMAAFKE 811
 QY 576 GFNELIPQHLAKPDKOLELILIGGLDKIDLNDKWNTRLKHCA-DSNIVRWFWQAVET 934

[6]

CHARACTERIZATION.

MEDLINE=98351563; PubMed=9698277;
Nuber U., Schwarz S.E., Schaffner M.;
"The ubiquitin-protein ligase b6-associated protein (b6-AP) serves as
its own substrate.";
Eur. J. Biochem. 254:643-649 (1998).

-!- FUNCTION: INTERACTS WITH THE E6 PROTEIN OF THE CANCER-ASSOCIATED
HUMAN PAPILLOMAVIRUS TYPES 16 AND 18. THE B6/E6-AP COMPLEX BINDS
TO AND TARGETS THE P53 TUMOR-SUPPRESSOR PROTEIN FOR UBIQUITIN-
MEDIATED PROTEOLYSIS. IT IS AN E3 UBIQUITIN-PROTEIN LIGASE WHICH
ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE
FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO
TARGETED SUBSTRATES. IT CAN TARGET ITSELF FOR UBIQUITINATION IN
VITRO AND EFFICIENTLY PROMOTES ITS OWN DEGRADATION IN VIVO. IT
APPEARS THAT ONLY UNMODIFIED E6-AP MOLECULES CAN BIND EFFICIENTLY
TO P53 IN THE PRESENCE OF THE HPV E6 ONCOPROTEIN.

-!- SUBCELLULAR LOCATION: Nuclear (Probable).

-!- ALTERNATIVE PRODUCTS:

-!- Event=Alternative splicing; Named isoforms=3;
Name=I1;
IsoId=Q05086-1; Sequence=Displayed;
Name=I;
IsoId=Q05086-2; Sequence=VSP_006705;
Name=I11;
IsoId=Q05086-3; Sequence=VSP_006706;

-!- DISEASE: DEFECTS IN UBES3 ARE A CAUSE OF ANGELMAN SYNDROME (AS),
ALSO KNOWN AS 'HAPPY PUPPET SYNDROME'. IT IS CHARACTERIZED BY
FEATURES OF SEVERE MOTOR AND INTELLECTUAL RETARDATION,
MICROCEPHALY, ATAXIA, FREQUENT JERKY LIMB MOVEMENTS AND FLAPPING
OF THE ARMS AND HANDS, HYPOTONIA, HYPERACTIVITY, HYPOPIGMENTATION
SEIZURES, ABSENCE OF SPEECH, FREQUENT SMILING AND EPISODES OF
PAROXYSMAL LAUGHTER, AND AN UNUSUAL FACIES CHARACTERIZED BY
MACROSTOMIA, A LARGE MANDIBLE AND OPEN-MOUTHED EXPRESSION, A GREAT
PREDISPOSITION FOR PROTRUDING THE TONGUE ('TONGUE THRUSTING'), AND AN
OCCIPITAL GROOVE.

-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
THIOLESTER FORMATION.

-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.

-!- SIMILARITY: SOME, IN THE C-TERMINAL HALF, TO RAT 100 kDa PROTEIN.

THIS SWISS-PROT entry is copyright. It is produced through a collabora-
tion between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no
modified and this statement is not removed. Usage by and for commer-
cial entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
or send an email to license@isb-sib.ch).

EMBL; X98021; CAA66653.1; --
EMBL; X98027; CAA66653.1; JOINED.
EMBL; X98022; CAA66653.1; JOINED.
EMBL; X98023; CAA66653.1; JOINED.
EMBL; X98024; CAA66653.1; JOINED.
EMBL; X98025; CAA66653.1; JOINED.
EMBL; X98026; CAA66653.1; JOINED.
EMBL; X98028; CAA66653.1; JOINED.
EMBL; X98029; CAA66653.1; JOINED.
EMBL; X98030; CAA66653.1; JOINED.
EMBL; X98033; CAA66656.1; --
EMBL; X98031; CAA66654.1; --
EMBL; X98032; CAA66655.1; --
EMBL; L07557; AAA35542.1; ALT_INIT.
EMBL; AF016708; AAB69154.1; --
EMBL; AF016703; AAB69154.1; JOINED.
EMBL; AF016704; AAB69154.1; JOINED.
EMBL; AF016705; AAB69154.1; JOINED.
EMBL; AF016706; AAB69154.1; JOINED.
EMBL; AF016707; AAB69154.1; JOINED.
EMBL; AF016708; AAB69154.1; JOINED.
EMBL; U84404; AAB49301.1; --

432 STDNLYVLQINPDSSINPDHLSYFHFVGRWGLAVFHHVGHVINGFTVPYKOLIGKPIQL 491
 593 D-ESTKLFWFNPSF---ETEGQTLIGVIGLAIYNNCLDVFHFVYRKMGKGTG 648
 492 SDLESVPDELKSLVWL--ENDITPVLDTFCV-EHNAFGRILQHELKPNRNVPVTEE 548
 649 RDLGSHFVLYQSLKDLLEVEGNVEDDMITFQISQDLFGNPMYDLKENGDKIPITNE 708
 549 NKKEVRLVYVNRFWRGIEQFLALQGF---NELIPQHLKPPFQKLELILGLDKI 604
 709 NRKGFVNYSYIILNKSVKEQKAFRRGFHMVTNESPILYFRP---ESTELLICGRNL 765
 605 DLNWKSNTRLK-HCVADSNIVRWFQAVETFEERRARLLQVGTGSTRVPLQGFALQ 663
 766 DFOALEETTEYDGGVTRDSVLIREFWEIVHSFTDEQKRLFLQTTGTDRAVGGGLKLM 825
 664 STGAAGPLFTIHLIDANTDNLKPAHCFENRDIPIPPYSEYKLYEKLITAVEETCGFV 722
 826 IIAKNGP-----DTERLPTSHTCFNVLLLPYSSKEKLERLLKAITYAKFGFM 874

3ULT 11
 1A MOUSE
 UE3A MOUSE STANDARD; PRT; 885 AA.
 008759; P97482;
 30-MAY-2000 (Rel. 39, Last sequence update)
 30-MAY-2000 (Rel. 39, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Ubiquitin-protein ligase E3A (EC 6.3.2.-) (Oncogenic protein-associated protein E6-AP).
 UBE3A.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=C57BL/6 X CBA;
 MEDLINE=97326076; PubMed=9182527;
 Hatakeyama S., Jensen J.P., Weisman A.M.;
 "Subcellular localization and ubiquitin-conjugating enzyme (E2) interactions of mammalian HECT family ubiquitin protein ligases.";
 J. Biol. Chem. 272:15085-15092 (1997).
 [2]
 SEQUENCE FROM N.A.
 STRAIN=BALB/c; TISSUE=Brain;
 MEDLINE=97264343; PubMed=9110176;
 Sutcliffe J.S., Jiang Y.-H., Galjaard R.-J., Matsuura T., Fang P., Kubota T., Christian S.L., Bressler J., Cattanch B., Ledbetter D.H., Beaudet A.L.;
 "The E6-AP ubiquitin-protein ligase (UBE3A) gene is localized within a narrowed Angelman syndrome critical region.";
 Genome Res. 7:368-377 (1997).
 -1- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES.
 -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 -1- TISSUE SPECIFICITY: MOST ABUNDANT IN BRAIN, HEART AND THYMUS.
 -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOESTER FORMATION.
 -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
 -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-38 IS THE INITIATOR.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

 EMBL; U96636; AAB63361.1; --

DR U82122; AAB47756.1; --
 DR MGD; MGI:105098; Ube3a.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0005834; C:nucleus; IDA.
 DR GO; GO:0005515; F:protein binding activity; IPI.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
 DR InterPro; IPR000569; HECT_domain.
 DR Pfam; PF00632; HECT; 1.
 DR SMART; SM00119; HECTC; 1.
 DR PROSITE; PS0237; HECT; 1.
 KW Nuclear protein; Ub1 conjugation pathway; Ligase.
 FT DOMAIN 405 410
 FT BINDING 853 885
 FT CONFLICT 187 187
 FT CONFLICT 301 301
 FT CONFLICT 343 343
 FT CONFLICT 384 384
 FT CONFLICT 460 460
 FT CONFLICT 535 535
 FT CONFLICT 577 582
 FT CONFLICT 584 584
 FT CONFLICT 587 588
 FT CONFLICT 597 597
 FT CONFLICT 623 627
 FT CONFLICT 725 726
 FT CONFLICT 817 817
 FT CONFLICT 869 870
 FT CONFLICT 885 AA; 101175 MW; 55D885E080CCB699 CRC64;
 SQ SEQUENCE 885 AA; 101175 MW; 55D885E080CCB699 CRC64;
 Query March 12.4%; Score 483; DB 1; Length 885;
 Best Local Similarity 28.5%; Pred. No. 4.4e-27;
 Matches 152; Conservative 103; Mismatches 216; Indels 62; Gaps 18;
 QY 209 VRGSLQTPQNRPHGHOS-PELPE-----GYEQRTVQGVYFLHTQTGVSTWHDPRIP 260
 DB 395 VGGVDVTHNNEDESPESSELTIQELLGERRNKGKGRVDPLETELGVKTL-DCRKP 453
 QY 261 RLNSVNCDELGLPGGWEVRVSVGRIVFVHNNTTOFTDPRLHHNHOCQLKPSQ 320
 DB 454 LISFEESINE--PLNDVLEMDKDYT--FFKVTENKNGSFMTCFFILNAVTKNLGYDNR 509
 QY 321 PLPLPSGSLDEELPAQRYERDLVKLRLHSLQLOPQAGHCRIVSREREIPEES-- 378
 DB 510 -IRMYSE-----RRITVL-YSLVQGOQLAPYLTLYKVRDHIIDALV 549
 QY 379 -YRQIMKVRPKDKKLMKFRGEBGLDYGVARWLYLLCHEMLNPPYGLFYQYSTDNIY 437
 DB 550 RLEMIAENPADLKKLYVEFEGEQVDEGVSKEFFQLVWVEIFNPDMFTYD-EATK 608
 QY 438 MLQINPDSSINPDHLSYFHFVGRIMGLAVFHHVGHVINGFTVPYKOLIGKPIQLSLESV 497
 DB 609 LFWNPSF---ETEGQTLIG-ILGLAIYNNCLDVFHFVYRKMGKGTGPRDLGDS 664
 QY 498 DPELHKSILVWL--ENDITPVLDTFCV-EHNAFGRILQHELKPNRNVPVTEENKVEYV 554
 DB 665 HPVLYQSLKDLLEVEGVEDDMITFQISQDLFGNPMYDLKENGDKIPITNEKKEFV 724
 QY 555 PLYVNRWFMGRIEAQLALQGF---NELIPQHLKPPFQKLELILGLDKILDNWK 610
 DB 725 ISYDDYILNKSVKEQKAFRRGFHMVTNESPILYFRP---ESTELLICGRNLDFQALE 781
 QY 611 SNTRLK-HCVADSNIVRWFQAVETFEERRARLLQVGTGSTRVPLQGFALQGGTGAAG 669
 DB 782 ETTEYDGGYTRSVVIREFWEIVHSFTDEQKRLFLFTTGTDRAPVGGGLKLMIAKNG 841
 QY 670 PRLTIHLIDANTDNLKPAHCFENRDIPIPPYSEYKLYEKLITAVEETCGFV 722
 DB 842 P-----DTERLPTSHTCFNVLLLPYSSKEKLERLLKAITYAKFGFM 884

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

EMBL; D25215; BAA04945.1; --
Genew; HGNC:4876; HERC3.
MIM; 605200; --
InterPro; IPR000569; HECT domain.
InterPro; IPR000408; Reg_Chrcondens.
Pfam; PF00632; HECT; 1.
Pfam; PF00415; RCC1; 4.
PRINTS; PR00633; RCCNDNSATION.
SMART; SM00119; HECT; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS00625; RCC1; 1; FALSE_NEG.
PROSITE; PS00626; RCC1; 2; 4.
PROSITE; PS00012; RCC1; 3; 7.
PROSITE; PS00012; RCC1; 3; 7.
Ubl conjugation pathway; Ubl conjugation; Repeat.
REPEAT 1 51
REPEAT 52 101
REPEAT 102 154
REPEAT 156 207
REPEAT 208 259
REPEAT 261 311
REPEAT 313 366
DOMAIN 951 1050
BINDING 1018 1018
SEQUENCE 1050 AA; 117188 MW; 5F08AIDE1P40B912 CRC64;
Query Match 11.6%; Score 449.5; DB 1; Length 1050;
Best Local Similarity 28.5%; Pred. No. 1.5e-24;
Matches 124; Conservative 86; Mismatches 184; Indels 41; Gaps 11;
311 HQCLK-RPS---QPLPSPSGSLEDELPQRYERLDVQLKVLKRLHLSLQ-----359
633 HQAGKARPSIIQDTVILCSYFFIFDAQKTMLOTDAELQXQVANGANLQNVFLLTL 692
360 ----POAGHRIEVSREIFESYRQIMKMRPKDKRLVMKFRGEGLDYGQVAREWLY 415
693 EPLLARSPLVLVRRNNLVGDALRELSIHSIDILKKPLKLVFDGSEAVDAGGVTKFEFL 752
416 LICHEMLNPYGLFOYSDNYMLQINPDSSINPDHLSYFHFVGRINGLAVFHGVINGG 475
753 LLLKELLPIYGMFTYYQDS-NLLWFSDTCFV--EH-NWPHLLGITCGLAINSTVVDLH 808
476 FTVPYKQLQKPIQSLDLESVDPELHKLWILE--NDITPVLDTFCVE----HNAP 528
809 PFLALYKLLNVKPGLEDKLSPTGSRSLQELLDYPGEDV---BETFCNFTICRESY 864
529 GRILQHELKPGNRPVPTVEENKKEYRLVYNNRPMRGIEAQFLAQGFNELIPQHLXP 588
865 GVEQKQLPGGDNTVYCKDRQRFVDAYVNVYQISVHEWYTAFFSGFLKVCQGVLEL 924
589 FQKLELIIGLDLIDNDKSNTRLK-HCVADSNIVRMFQAVETFDERRARLLQFV 647
925 FQPSLRAMVGNVNNWEELEEARAIYKGYSAHTPTVKLFWETTFHFFLEKXKLLFL 984
648 TGSTRVPGFVALQSGTAGAPRLFTTHLDANTDMLPKATCFNRIDIPPPYSEKLY 707
985 TGSDRIPYGMASLQ-----IVIQSTASGEVLPVAHTCYNLLDLPKYSKREILS 1034
708 EKLLTAVEETCGFAV 722
1035 ARLTQALDNYEGFSL 1049

```

SULT 14
IB_HUMAN
_TRIB_HUMAN
Q14669; Q15644;
STANDARD;
PRT; 1992 AA.

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Thyroid receptor interacting protein 12 (TRIP12).
GN TRIP12 OR KIAA0045.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
RN [2]
RP SEQUENCE OF 1801-1992 FROM N.A.
RX MEDLINE=95295737; PubMed=7776974;
RA Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RT "Two classes of proteins dependent on either the presence or absence
RT of thyroid hormone for interaction with the thyroid hormone
RL receptor.";
RL Mol. Endocrinol. 9:243-254(1995).
CC -!- FUNCTION: COMPONENT OF P4700, AN ATP-DEPENDENT MULTISUBUNIT
CC PROTEIN THAT ACTIVATES THE PROTEOLYTIC ACTIVITIES OF THE
CC MULTIFUNCTIONAL PROTEINASE (20S PROTEASOME) OF THE 26S COMPLEX.
CC SPECIFICALLY INTERACTS WITH THE LIGAND BINDING DOMAIN OF THE
CC THYROID HORMONE RECEPTOR (IN A THYROID HORMONE T3-INDEPENDENT
CC MANNER) AND WITH RETINOID X RECEPTOR (RXR). COULD BE E3 UBIQUITIN-
CC PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-
CC CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND THEN DIRECTLY
CC TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES.
CC -!- SUBUNIT: P4700 IS COMPOSED OF AT LEAST 16 DISTINCT PEPTIDES
CC RANGING IN MOL. WT. FROM 20-112 kDa.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Query Match 9.2%; Score 358; DB 1; Length 1992;
Best Local Similarity 25.1%; Pred. No. 1.7e-17;

Matches 142; Conservative 94; Mismatches 208; Indels 122; Gaps 24;

241 QVYFLHTQGVSTWHDPRIP-----RDNSVN-----CDELGPLPGHVEVSTV 284
 1451 EVYLIPTPPENITFDPSLDVILLRLVLAHRAIKRYWYLYVDNAMCKEI--IPTSFINSKL 1508
 285 SGRIYFYDHNRTTQTDRLRHIMHQCQLKEPQPLPSEGSLEDEELPAQRYERD- 343
 1509 TAKA-----NR--QLQDPLVMTGNIPTWLTGLKTCPPFPPTDQMLFVYTAQRDR 1560
 344 LVQKLKVLRLHEL-----SLOQPQACHCRIEVSREIEPEESYRQIMKWRKDL-KRRLM 395
 1561 AMQRLDITNEINQSDQSRVAPRDRKRTVNRRELKQA-ESVM-----QDLGSSRAM 1615
 396 VKFRGEGGLDYG-GVAREWLYLLCHEMLNPFYGLF-----QYSTDNIYMLQ- 440
 1616 LEIQYENVEGTGLPTLEFVALVSQELQADILGLWRGEBVTLSNPKGSQEGTKYIQNLQ 1675
 441 ---INPDSSINPHLS-----YHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGRPIQLS- 492
 1676 LPAFPRTAKPAHIAKVKRERFLGKLMAKAIMDFRLVDLPLGLFFYKWLMLRQETSLTS 1735
 493 -DLESVDPELHKLWMI-----LENDITPV-----LDHT 520
 1736 HDLFDIDPVARSVYHLEDIVRQKKEQDKSQTKESLOYALETITMNGCSVEDLGLDFT 1795
 521 FVEHNAFGRILQHELKPGNRVPTVEENKKYRVLVYVNRFRMGIEAQFALQKGFNEL 580
 1796 L-----POGFNI-----ELKGGKQIPVTIHNLEBYRLVIFWALNEGVSQFDSFRDPESV 1848
 581 IPOHLKLPFQKQLELIIGSLDKIDLNDKSNLKHG-----VADSNIVRWFQAVET 634
 1849 FPLSHLOYFPELQQLCG-SKAD--TWDAKILMECCRPDGHYTHBRVKFIFEILSS 1905
 635 FDEERARLLQFVTGSTRVPLOFGKALQGSTGAAGPRLFTIHLIDANTDN-----LPKAHT 690
 1906 FDNEQQLFLQFVTGSTRVPLVGGFRSL-----NPLTIIVAKTPESTENPDDFLPVSMT 1958
 691 CFNRIIDIPPESEYKLYEKLITAVEE 716
 1959 CVNYLKLPDYSSIEIMREKLLIAARE 1984

GENE NAME AND GENE DISRUPTION.
 MEDLINE=9077972; PubMed=9858558;
 Wang G., Yang J., Huibregtse J.W.;
 "Functional domains of the reps ubiquitin-protein ligase."
 Mol. Cell. Biol. 19:342-352(1999).
 -!- FUNCTION: PROBABLE E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS
 UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A

CC THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED
 CC SUBSTRATES. NON ESSENTIAL
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X99960; CAA68221.1; -;
 CC EMBL; Z72663; CAA96853.1; -;
 CC FIR; S64155; S64155.
 CC SGD; S0003109; YGL141W.
 CC InterPro; IPR000569; HECT_domain.
 CC Pfam; PF00632; HECT; 1.
 CC SMART; SM00119; HECT; 1.
 CC PROSITE; PS50237; HECT; 1.
 CC Ub1 conjugation pathway; Ligase.
 CC DOMAIN 810 810
 CC BINDING 878 878
 CC FT SEQUENCE 910 AA; 105565 MW; 8DB932A48A91CD19 CRC64;
 CC SQ
 Query Match 9.1%; Score 354; DB 1; Length 910;
 Best Local Similarity 23.7%; Pred. No. 1.1e-17;
 Matches 136; Conservative 93; Mismatches 212; Indels 132; Gaps 20;

QY 220 PFGHOSPELPEG-----YEQRTVQGVYFLTQTGVSTWHDPRIPRLNSVNCDELG 272
 Db 395 PKSERNPMLKAVPLLSKYER-----DSRLHFLSTENN----- 428
 QY 273 PLPPGWEVRSTVSGRIYFVDHNNRTQTDRLRHIMN-----HOCOLKEPSQPLPSE 327
 Db 429 ---PTWE-----NSEKQFLNRFEELOEYEDLYREHLEESDE-DMEKE 470
 QY 328 GSLEDEELPAQ-----RYERDLVQKLKVLRLHLSLOQP----- 360
 Db 471 IDLKERPLKSLLLKMKRLKSLRPRKLEILLELPFFPIPFPERVDFLYMFIALDKR 530
 QY 361 ---QACHCRIE-----VSREEIEESYRQIMKMRPK---DLKRLAVKF 398
 Db 531 LSLDDHNLINFTWASTGMRKQSAIISRDVLEDAFNAPNSIGERFKASLDVTFINEF 590
 QY 399 RGEGLDYGAVAREWLYLLCHE-MLNPYVGLFYQYSTDNIMYMLQINPDSSINPDLSYFHF 457
 Db 591 GEEAGIDGGITKEFLTTSVDSBGFDPKHLEFR--TNDRY--ELYPSVWYDATKLYWF 646
 QY 458 VGRIMGLAVFHGHYINGGFTVPFYKQLL-----GKPIQLSDLESVDPELHKLWILE--N 511
 Db 647 LQKVGKCLYEHVLIDVDFADFLKLLNYSNGFLSFSGLSYDSLYNNLIKLANNTT 706
 QY 512 DITPVLDTHTFCV-EHNAFCRILQHELKPGNRVPTVEENKKEYVRLVYVNRFRMGIEAQF 570
 Db 707 DAIKSLDLTFEIDEPESAKVV--DLIPNGSKTYTKDNLVLYTKVTDYKLNKRCFKPV 764
 QY 571 LALQKGFNELIPQHLKPFDOKELELIIGG-LDKIDLNDKSNTRLKHCVADSNIVRWF 629
 Db 765 SAFHGGLSVIAAPHWMENFNSIELQMLISGERDIDLDKLSNTEYGGYKEEDQTIYDFW 824
 QY 630 QAVETFDERRARLLQFVTGSTRVPLOFGKALQGSTGAAGPRLFTIHLIDANTD--NLPK 687
 Db 825 EVLNBEKFEKLNFLKFTVSVPQAPLOQFALDKPFG-----IRNAGTEKYLPT 874
 QY 688 AUTCPNRIDIPPESEYKLYEKLITAVEETCGF 720
 Db 875 ASTCVNLLKLPDYNKTLIRKLLYAINSGARF 907

arch completed: February 20, 2004, 15:26:05
o time : 13.8131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 protein - protein search, using sw model
in on: February 20, 2004, 15:25:26 ; Search time 33.9137 Seconds
(without alignments)
5501.382 Million cell updates/sec

tle: US-10-009-945-2
irect score: 3884
equence: 1 GGSIKIRLAVLCANLAKK.....EKLYEKLITAVETCGFAVE 723

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 830525 seqs, 258052604 residues

tal number of hits satisfying chosen parameters: 830525

imum DB seq length: 0
imum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : SPTREMBL 23:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3792.5	97.6	728	11 Q8K300	Q8K300 mus musculus
2	2206	56.8	1061	5 Q9V853	Q9V853 drosophila
3	1895	48.8	355	11 Q8BSC0	Q8BSC0 mus musculus
4	1559.5	40.2	295	4 Q8NDD8	Q8NDD8 homo sapien
5	1341	34.5	786	3 O14326	O14326 schizosach
6	1340.5	34.5	288	4 Q96DE7	Q96DE7 homo sapien
7	1318	33.9	854	11 O54971	O54971 mus musculus
8	1301	33.5	862	4 Q9BY75	Q9BY75 homo sapien
9	1288.5	33.2	903	4 Q96F66	Q96F66 homo sapien
10	1269	32.7	911	4 Q8NSA7	Q8NSA7 homo sapien
11	1247	32.1	955	4 Q96PU5	Q96PU5 homo sapien
12	1245.5	32.1	703	4 Q9H451	Q9H451 homo sapien
13	1245.5	32.0	739	4 Q43584	Q43584 homo sapien
14	1241.5	32.0	820	4 Q9NT88	Q9NT88 homo sapien
15	1239	31.9	922	4 Q9H0M0	Q9H0M0 homo sapien
16	1239	31.9	949	5 Q9Y0H4	Q9Y0H4 drosophila

17	1234	31.8	870	11 Q9DBH0	Q9DBH0 mus musculus
18	1228.5	31.6	855	4 Q8WU09	Q8WU09 homo sapien
19	1228.5	31.6	858	4 Q9BW58	Q9BW58 homo sapien
20	1228.5	31.6	971	13 Q42573	Q42573 xenopus lae
21	1228.5	31.6	995	4 Q43165	Q43165 homo sapien
22	1227.5	31.6	887	11 Q62940	Q62940 rattus norv
23	1226	31.6	854	4 Q9H2M4	Q9H2M4 homo sapien
24	1225	31.5	870	4 Q96CZ2	Q96CZ2 homo sapien
25	1225	31.5	870	4 Q96CZ2	Q96CZ2 homo sapien
26	1225	31.5	887	11 Q8BGB3	Q8BGB3 mus musculus
27	1224	31.5	855	11 Q8CF10	Q8CF10 mus musculus
28	1223.5	31.5	835	11 Q8BRT9	Q8BRT9 mus musculus
29	1215	31.3	855	11 Q99PK2	Q99PK2 mus musculus
30	1214.5	31.3	834	5 Q9SR64	Q9SR64 drosophila
31	1213	31.2	794	5 Q9NZ27	Q9NZ27 caenorhabdi
32	1210	31.2	792	5 Q9SXU3	Q9SXU3 caenorhabdi
33	1207.5	31.1	376	13 Q98T84	Q98T84 xenopus lae
34	1200.5	30.9	258	11 Q9CSE3	Q9CSE3 mus musculus
35	1196.5	30.8	617	11 Q8BZ23	Q8BZ23 mus musculus
36	1196.5	30.8	956	5 Q8TOR6	Q8TOR6 drosophila
37	1195.5	30.8	518	5 Q8T0C8	Q8T0C8 drosophila
38	1191	30.7	1007	5 Q9VVI3	Q9VVI3 drosophila
39	1189	30.6	777	11 Q88758	Q88758 mus musculus
40	1151	29.6	838	5 Q9STQ0	Q9STQ0 drosophila
41	1129	29.1	724	5 Q9BKW4	Q9BKW4 caenorhabdi
42	1047	27.0	684	4 Q00307	Q00307 homo sapien
43	1026.5	26.4	671	3 Q9UTG2	Q9UTG2 schizosach
44	1002	25.8	1585	4 Q9HCC7	Q9HCC7 homo sapien
45	991	25.5	759	11 Q9C863	Q9C863 mus musculus

ALIGNMENTS

RESULT 1

Q8K300 PRELIMINARY; PRT; 728 AA.

ID Q8K300

AC Q8K300; 22, Created)

DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)

DE Similar to E3 ubiquitin ligase SMURF1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RP [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

DR EMBL; BC029097; AAH29097.1; -.

DR InterPro; IPR000008; C2.

DR InterPro; IPR000569; HECT domain.

DR InterPro; IPR001202; WW_Rsp5_WWP.

DR Pfam; PF00168; C2_1.

DR Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 2.

DR SMART; SM00239; C2; 1.

DR SMART; SM00119; HECTC; 1.

DR SMART; SM00456; WW; 2.

DR PROSITE; PS00499; C2 DOMAIN_1; 1.

DR PROSITE; PS00004; C2_DOMAIN_2; 1.

DR PROSITE; PS00237; HECT; 1.

DR PROSITE; PS01159; WW DOMAIN_1; 1.

DR PROSITE; PS00020; WW_DOMAIN_2; 2.

DR Ligase.

SQ SEQUENCE 728 AA; 83083 MW; C355291B9D8AD757 CRC64;

Query Match 97.6%; Score 3792.5; DB 11; Length 728;
Best Local Similarity 98.1%; Pred. No. 2e-303;
Matches 708; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

2 GSSIKILTLVLCANLAKKDFEFLPDPFAKIVVDGSGQCHSTDTKNTLDPKWNQHYDLY 61
10 GSSIKILTLVLCANLAKKDFEFLPDPFAKIVVDGSGQCHSTDTKNTLDPKWNQHYDLY 69
62 VGKTDSTISVWNHKKIKKQAGAGFLGCVRLLSNAISRLKOTGYORLDCIKLPNSDTPAV 121
70 VGKTDSTISVWNHKKIKKQAGAGFLGCVRLLSNAISRLKOTGYORLDCIKLPNSDTPAV 129
122 RGQIVLSLQTRDRLTGTSVVDRCRGLLENGTVYDSGGRPLSCFMEPEAPYDSTGAA 181
130 RGQIVLSLQTRDRLTGTSVVDRCRGLLENGTVYDSGGRPLSCFMEPEAPYDSTGAA 189
182 AGGNCRFVSPSQDQRLQALRNPDRVLSGLTPQNRPHGQSPELPEGYEQRITVQGG 241
190 AGGNCRFVSPSQDQRLQALRNPDRVLSGLTPQNRPHGQSPELPEGYEQRITVQGG 249
242 VYFLHTGTGTVSTHWDPRIPDLNSVNCDELGPLPGWEVRSVTSGRIFVVDNNRTTFT 301
250 VYFLHTGTGTVSTHWDPRIPDLNSVNCDELGPLPGWEVRSVTSGRIFVVDNNRTTFT 309
302 DPLRLHIMNHQCOKEPSQPLPSPSGSLEDELPQRYERDLVQKLKVLRLHLSLQOPQ 361
310 DPLRLHIMNHQCOKEPSQPLPSPSGSLEDELPQRYERDLVQKLKVLRLHLSLQOPQ 369
362 AGHCRISVRSREIPEESYRQIMKWRPKDLKRLMWKFRGEBGLDYGVARWLYLLCHEM 421
370 AGHCRISVRSREIPEESYRQIMKWRPKDLKRLMWKFRGEBGLDYGVARWLYLLCHEM 429
422 LNPYVYGFQVSTQNIYMLQNPSSINPHLSYFHFVGRIMGLVPHGHYINGGTPVPFY 481
430 LNPYVYGFQVSTQNIYMLQNPSSINPHLSYFHFVGRIMGLVPHGHYINGGTPVPFY 489
482 KOLLGKPIQLSDLESVDPELHKLVLNENDITPVLVDHTFCVHEHNAFRILOHELKPNGR 541
490 KOLLGKPIQLSDLESVDPELHKLVLNENDITPVLVDHTFCVHEHNAFRILOHELKPNGR 549
542 NVPVTENKKEYVRLVYVNRVFMGIEBAQFLALQKGFNELIPQHLKPKPDQKELELIIIGL 601
550 NVPVTENKKEYVRLVYVNRVFMGIEBAQFLALQKGFNELIPQHLKPKPDQKELELIIIGL 609
602 DKIDLNDKNTLKHCVADSNIVRWFQAVETFDREARLLQFVTGSTRVPLQGFAL 661
610 DKIDLNDKNTLKHCVADSNIVRWFQAVETFDREARLLQFVTGSTRVPLQGFAL 669
662 QGSGTAGPRLFTIHLIDANTDNLKPAHCFNRIDIPPYESYEKLYEKLITAVETCGFA 721
670 Q---GAAGPRLFTIHLIDANTDNLKPAHCFNRIDIPPYESYEKLYEKLITAVETCGFA 726
722 VE 723
727 VE 728

SULT 2

V853

Q9V853 PRELIMINARY; PRT; 1061 AA.

Q9V853; Q9U3W2;

01-MAY-2000 (T-EMBLrel. 13, Created)

01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

01-MAR-2003 (T-EMBLrel. 23, Last annotation update)

CG4943 protein (E3 ubiquitin ligase) (ubiquitin-protein ligase) (Smad-

ubiquitin E3 ligase Smurf1).

LACK OR SMURF OR SMURF1 OR CG4943.

Drosophila melanogaster (Fruit fly).

Eukaryota; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Echydnoidea; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

[1]

SEQUENCE FROM N.A.

STRAIN=Berkley;

MEDLINE=20196006; PubMed=107311132;

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Anil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong P., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Furi V., Reese M.G.,
RA Rainart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster";
RA Science 287:2185-2195(2000).
RL [2]
RL SEQUENCE FROM N.A.
RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Genzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
RA "Sequencing of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP Misra S., Crosby M.A., Matthews B.B., Rayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Scari S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RP FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```

1 [6]
2 SEQUENCE FROM N.A.
3 Laurence A., Hawley S.;
4 "Molecular cloning of a type E3 Ubiquitin ligase.";
5 Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
6 [7]
7 SEQUENCE FROM N.A.
8 Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;
9 "The DSmurf ubiquitin-protein ligase restricts BMP signaling spatially
10 and temporally during Drosophila development.";
11 Dev. Cell 1:0-0 (2001).
12 [8]
13 SEQUENCE FROM N.A.
14 Liang Y.-Y., Lin X., Feng X.-H.;
15 "dSmurf, a Smad-ubiquitin E3 ligase, specifically targets dpp-
16 activated Mad protein for degradation.";
17 Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
18 -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
19 EMBL; AF003802; AAF57824.3; -
20 EMBL; AF216521; AAF21125.1; -
21 EMBL; AF416571; AAL09691.1; -
22 EMBL; AF464851; AAMC9646.1; -
23 HSSP; Q1326; 1PIN.
24 Flybase; FBgn029006; lack.
25 InterPro; IPR000008; C2.
26 InterPro; IPR000569; HECT_domain.
27 InterPro; IPR002349; WW.
28 InterPro; IPR001202; WW_Reps_WWP.
29 Pfam; PF00168; C2; 1.
30 Pfam; PF00397; WW; 3.
31 PRINTS; PR00403; WWDOMAIN.
32 SMART; SMC0239; C2; 1.
33 SMART; SMC0456; WW; 3.
34 PROSITE; PS00499; C2_DOMAIN_1; 1.
35 PROSITE; PS00004; C2_DOMAIN_2; 1.
36 PROSITE; PS0237; HECT; 1.
37 PROSITE; PS01159; WW_DOMAIN_1; 2.
38 PROSITE; PS00020; WW_DOMAIN_2; 3.
39 Ligase.
40 SEQUENCE 1061 AA; 115675 MW; 68BCC550F5129163 CRC64;
41 Query Match 56.8%; Score 2206; DB 5; Length 1061;
42 Best Local Similarity 43.08; Pred. No. 1.8e-172;
43 Matches 454; Conservative 108; Mismatches 148; Indels 346; Gaps 14;
44
45 6 KIRTVLCANLAKQFRLPDPFAKIVVDGSGQCHSTDTVXNTLDPKNNQHYDLYVGT 65
46 14 KVRITICARNLARKOLFRLPDPFAKIVVDGSGQCHSTDTVXNTLDPKNNQHYDLYVGT 73
47
48 66 DSITISVNNHKKTHKKGAGFLGCVRLLSNAISRLKDTGYQRLDLCKLNPSTDAVRQI 125
49 74 DAITIVNQRKHK--GSGFLGCVRIAPAFNIQSLKGAGFQRLDLGKUSPDDDLVRQI 131
50
51 126 VWSLQTRDRIGTG-----GSVVDCRGLLENE-----GTIVY----- 155
52 132 IISLLSKDGPSSNPLAIVGSDVRGSEDDSDSLPEGWEERTDNGRYVYVNHATK 191
53 156 ----- 155
54
55 192 STQMDRPRQPGVGVSSHATSPQORHNTNGSGDRQAPAGTRSTCTNLMMNGHRSRL 251
56 156 -----EDSGRPLSCFMEEPAPYTDSTGAAGGNCRFVE----- 191
57 252 SVTASDERRHSTELISVGKENTSPITPVSATTPFGKTSNNSAGG--RTLEQRPT 308
58 192 -----SPQDQRLQARLNDVKGSLQTONRPHGQSPE----- 227
59 309 NEPATPSTTSASVRLNSD--NHVKTPKQTNCHAPPSTPTSGTQQNVYNGNAQNG 366
60 228 ----- 227
61 367 STSNGSGQAQPCASNGHWGTQEDAAITTSPTTSPTRPHSGSPPTPNISPPASVTPSAN 426

```

```

228 ----- 227
427 GNVHSPNANSTPAGSGGSRSYTAATPQSRQRRSSROQGESESTRRRSSRGTRNGTSG 486
228 -----LPGYBORTTVQGVYFLHTQTGVSTWHPRI 259
487 GGGGGSGQRYASAAIAAQAARPFDLPPGYEWRTTQQGVFYHIPGVSTWHPRI 546
260 PDLNS--VNCDELGPLPGWEVSTVSGRIYFVDHNNRTTQFTDPLHHIM----- 309
547 PRDFDTQHLTDAIGFPLSGWEQRKTASGRVYFVDHNNRTTQFTDPLSGSILQWIRRG 606
310 ---NHQCQLKEPSQPLPLPS-----EG 328
607 VPTTSANAGTAPPSATPAPSAVAAPPOATPASNATPTTLTTNPPHRIYVDLPQG 666
329 SLEDELPQARYERDLVQKLKVLRLHLSLQOPQAGHCHREIVSREEIFEESYROIWKRPK 388
667 LLEGADL-LPKYRRDLVGLKRALRTELQTMQPSGHCHREIVSREEIFEESYRLIMKRAK 725
389 DLKRLMKFERGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNIMYMLQINPDASIN 448
726 DMKRLMKFERGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNIMYMLQINPDASIN 785
449 PDHLSYFHFVGRIMGLAVFHHYINGVTFVFFYKQLGKPIQLSLDESVDPELHKSIVWI 508
786 PDHLSYFHFVGRITGLIAVFGHCLDGGFTTFFYKQLLNKPIITLGDIEGVDPLHSLTWM 845
509 LENDITPVLDTFCVEHNAFGRILQHELKPNRNVVPTTEENKKEYVRLYNWRPMRGIEA 568
846 LESNISGIIESTFSVENNSFGALVVHELKPGGASIPVTEENKREYVKLYNRFMRGIEQ 905
569 QFLALQKGFNELIPOHLLKPPFOKELELIIGLDKIDLNDWKSNTLKLKCHVADSNIVRWF 628
906 QFLALQKGFCELLPSHLLRPDERELELVIGGISIDVNDWRNNTLKLKCHCTNETQVLP 965
629 WQAVETPDERARLLQFVTGSTRVLPQKALOGSTGAAGPRLFTIHL-IDANTDMLPK 687
966 WQVYESYSSEMARLLQFVTGSSRVLPQKALOGSTGAAGPRLFTIHLTADVPTQNLPK 1025
688 AHTCFNRIDIPYSEYKELTLTAVEETCGFAVE 723
1026 AHTCFNRIDIPYETVQLLCKLTOAVEETCGFAVE 1061
RESULT 3
Q8BSC0 PRELIMINARY; PRT; 355 AA.
AC Q8BSC0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical HECT domain (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK034736; BAC28813.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;
Query Match 48.8%; Score 1895; DB 11; Length 355;
Best Local Similarity 99.7%; Pred. No. 1.4e-147;

```

Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

369 VSREBIFESYQIMKMPKOLKRLMWKFRGEGLDYGGVAREWLYLLCHEMLNPFYGL 428
1 VSREBIFESYQIMKMPKOLKRLMWKFRGEGLDYGGVAREWLYLLCHEMLNPFYGL 60

429 FOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHYINGGFTVPFYKQLLGP 488
61 FOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHYINGGFTVPFYKQLLGP 120

489 IQLSDLESVDPELHLSLVWILENDITPVLDTFCVHNHAFGRILQHELKPNGRNVPTEE 548
121 IQLSDLESVDPELHLSLVWILENDITPVLDTFCVHNHAFGRILQHELKPNGRNVPTEE 180

549 NKKEYVRLYVNWRFMRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDLND 608
181 NKKEYVRLYVNWRFMRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDLND 240

609 WKSNTRELKHCVDASINVRWFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTGAA 668
241 WKSNTRELKHCVDASINVRWFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTGAA 300

669 GPRLFTIHLIDANTDNLPKAHTCFNRIDIPPPYESYEKLYEKLITAVEETCGFAVE 723
301 GPRLFTIHLIDANTDNLPKAHTCFNRIDIPPPYESYEKLYEKLITAVEETCGFAVE 355

SULT 4
NDD8 QNDD8 PRELIMINARY; PRT; 295 AA.

QNDD8;
01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein (Fragment).
DKFZF564H223.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain;
Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.,
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AL834242; CAD38919.1; -
InterPro; IPR000569; HECT_domain.
Pfam; PF00632; HECT; 1.
SMART; SM00119; HECTC; 1.
PROSITE; PS50237; HECT; 1.
Hypothetical protein.
NON_TER
1
SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;

Query Match 40.2%; Score 1559.5; DB 4; Length 295;
Best Local Similarity 98.7%; Pred. No. 4.6e-120;
Matches 293; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

427 GLFOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHYINGGFTVPFYKQLL 486
2 GLFOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHYINGGFTVPFYKQLL 61

487 KPIQLSDLESVDPELHLSLVWILENDITPVLDTFCVHNHAFGRILQHELKPNGRNVPVT 546
62 KPIQLSDLESVDPELHLSLVWILENDITPVLDTFCVHNHAFGRILQHELKPNGRNVPVT 121

547 EENKXEYRLYVNWRFMRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDL 606
122 EENKXEYRLYVNWRFMRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDL 181

607 NDWKSNTRELKHCVDASINVRWFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTG 666
182 NDWKSNTRELKHCVDASINVRWFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTG 723

QY 667 AAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPPYESYEKLYEKLITAVEETCGFAVE 723
Db 239 AAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPPYESYEKLYEKLITAVEETCGFAVE 295

RESULT 5
O14326 PRELIMINARY; PRT; 786 AA.
AC O14326;
DT 01-JAN-1999 (TrEMBLrel. 09, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ubiquitin-protein ligase.
GN SPBC1659.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]_TaxID=4896;
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Volckaert G., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; 299759; CAB16903.1; -
DR HSP; Q13526; 1P1N
DR GeneDB_Spombe; SPBC1659.11C; -
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PRO0403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
KW Ligase.
SQ SEQUENCE 786 AA; 89259 MW; 57B7A859F5497B9A CRC64;

Query Match 34.5%; Score 1341; DB 3; Length 786;
Best Local Similarity 36.6%; Pred. No. 2.1e-101;
Matches 304; Conservative 128; Mismatches 236; Indels 162; Gaps 21;

QY 6 KIRLTVLCANLAKKDFRLLPDPFAKIVVDGSGQCHSTDTVKNLDPKWNQHYDLYVGKT 65
Db 7 RVRFYVAAADGLSKRDLFRQDPDFAILTVDGE-QTHTTKVIKSNVFNWNEGEVTVKPS 65

QY 66 DSITTSVWNHKKHKQAGFLGCVRLLSNAISLKDYGQR-----LDLCK-LNPSDID- 119
Db 66 SVISRLFDQKKF-KKQDQGLGVSF-----RMREVGSFRSNREVSLRPLKKSSTN 118

QY 120 -AVRQIIVV-----SL 129
Db 119 LSVLGNLVKVPKIRAPAGNHSSTTANRTTPTTTTARTTTRTPTATTNTSNOST 178

QY 130 QTRDRIGTGGVDCRGLENNEGVYEDSGFGRPL-----SCWEEPA-----PYTDS 177
Db 179 SNSTNGTSAATNGTGTGAGTGASHRSS-----PVTNRQTNNTSALSNNAHMSFEDQ 234

QY 178 TGAAGG-----GNCRFVE-----SPSQDQRLQAQRLNPDV 209
Db 235 YGRLPFGVERADSLGRYYVDHNTRTTTWTWRASSINPVHNTSSDSQRLNHQNLHPD- 293

QY 210 RGSLOTQPNRPHGHQSP---ELPEGYQRTVQGVYFLHTGTGVSTWHDPRIP-----260

```

294 -----DSNPSLMQSDSGNDLFPGEWRYTDTGRPYFVDHNRITTTWDRNPLVRNG 346
261 -----RDLNSVNCDELGPPLPGWEYRSTVSGRIYFVDHNRITTTQFDRHLHIMHOCQ 314
347 GSSTVSLMOPQSLHGLPLSGWEMRLTNSARVYFVDHNTTTTWDOPRL----- 397
315 LKPSQPLPLPSGSLDEELPAQRVERDLVQKLVLRHLSLQOPOAGHCEIYSREEL 374
398 -----PSALDQVP--QYKDFRRLKIYFRSQPM-RPLPQCNVKNRRDHI 441
375 FEBSYQIMQRPKDLKRLMWKFRGEGLDGGVAREWLKYLHEMLNPNYGLFOXSTD 434
442 FEDSYABINRYSADHLKRLMIRFDEGEDGLDYGSLREFFLSHKMPDPIYCLFEYSV 501
435 NIYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGTFVPFYKQLLQKIQISDL 494
502 DNTLQINPHSSINPBLNYPFPIGVIGLAFHFRFLDAFFVWSLYKKLKRKYSLADM 561
495 ESVDPELHKSVMILENDITPVLDDHTFCVHNFAFRILQHELKPNRNVPTVEENKKEYV 554
562 ESIDAEFRSLKWLNDITGILDLTFSVEEDHFGVVRTVELITNGENIEVTENKKEYV 621
555 RLYNWRFMGIEAQLALQKGFNELIPOLLKPPDQKELELIIGGLDKIDLNDWKSNT 614
622 DLYTEWRSKRVRQQNFAYSGFVELVSPDLVNVFDERELELLIGGISDVVDWKSHT 681
615 LKHCVDASNIVRFWQAVTFDEERRARLLQFVTGSTRVPLQGFKALQSGAAGPRLF 674
682 YRYIATDPVKNFWIAGWKNRDSKLLQATGTSRIPVNGFRDLQSD--GPRKFT 738
675 IHLIDANT-DNLPKANTCNRIDIPPEYSEKLYEKLITAVEETCGFAVE 723
739 IE--KAGTDPQLPVAHTCNRLDLPYPSKDTLHEKLSLAVENTVGFNGE 786

SULT 6
6DE7
Q36DE7 PRELIMINARY; PRT; 288 AA.
Q36DE7;
01-DEC-2001 (T-EMBLrel. 19, Created)
01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
Similar to E3 ubiquitin ligase SMURF2 (Fragment).
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Lung;
Strausberg R.;
Submitted (JUN-2001) to the EMBL/GenBank/DBSJ databases.
EMBL; BC009527; AAH09527.1; -.
InterPro; IPR000569; HECT_domain.
Pfam; PF00632; HECT; 1.
SMART; SM00119; HECTC; 1.
PROSITE; PS50237; HECT; 1.
Ligase.
NON_TER 1
SEQUENCE 288 AA; 33255 MW; FE2B43E300DE66537 CRC64;
Query Match 34.5%; Score 1340.5; DB 4; Length 288;
Best Local Similarity 84.8%; Pred. No. 4.9e-102;
Matches 246; Conservative 23; Mismatches 18; Indels 3; Gaps 1;

434 DNTYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGTFVPFYKQLLQKIQISD 493
2 DDTYTLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGTFVPFYKQLLQKIQISD 61
494 LESVDPELHKSVMILENDITPVLDDHTFCVHNFAFRILQHELKPNRNVPTVEENKKEY 553
62 MELVDPELHKSVMILENDITGILDLTFSVEEDHFGVVRTVELITNGENIEVTENKKEY 121

```

```

QY 554 VRLYVNRFMGIEAQLALQKGFNELIPOLLKPPDQKELELIIGGLDKIDLNDWKSNT 613
DB 122 VRLYVNRFMGIEAQLALQKGFNEVIPOLLKPPDQKELELIIGGLDKIDLNDWKSNT 181
QY 614 RUKHCVDASNIVRFWQAVTFDEERRARLLQFVTGSTRVPLQGFKALQSGAAGPRLF 673
DB 182 RUKHCVDASNIVRFWQAVTFDEERRARLLQFVTGSTRVPLQGFKALQSGAAGPRLF 238
QY 674 TLIHLIDANT-DNLPKANTCNRIDIPPEYSEKLYEKLITAVEETCGFAVE 723
DB 239 TLIHLIDANT-DNLPKANTCNRIDIPPEYSEKLYEKLITAVEETCGFAVE 288

RESULT 7
CS4971 PRELIMINARY; PRT; 854 AA.
AC 054971;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Ubiquitin protein ligase.
GN ITCH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HeJ; TISSUE=Kidney;
RX MEDLINE=98122574; PubMed=9462742;
RA Perry W.L., Hustad C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,
RA Copeland N.G.;
RT "The itchy locus encodes a novel ubiquitin protein ligase that is
RT disrupted in al8H mice";
RL Nat. Genet. 18:143-146(1998).
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AF037454; AAB99764.1; -.
DR HSSP; Q13526; 1PIN.
DR MGD; MGI:1202301; Itch.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
KW Ligase.
SQ SEQUENCE 854 AA; 97975 MW; AA832D7BF0B9B3D8 CRC64;
Query Match 33.9%; Score 1318; DB 11; Length 854;
Best Local Similarity 35.4%; Pred. No. 1.9e-99;
Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

QY 6 KIRLVLCAL-NLAKKOFRLPDPFAKIVVGGSGCHSTDTVKNLTDPKNOHYDLYVGK 64
DB 9 QCIITVISAUKENKKNWFG-PSPIVETVD--GQSKKTEKKNNTSPRWKQPLTIVTVP 65
QY 65 TDSITISVYNNHKIKKQAGFLGCVRL-----LSNAISRLKDTGYQR----- 107
DB 66 TSKLCFRYWSHQL--KSDVLLGLTAGLDIYETLKSNNMKLEEVWMTLQVGGKEPTM 122
QY 108 --LDLCKLNPDSTAVRGOIVS-----LQTRD--RIGTCGS---VVDGR 145
DB 123 GDLSYC----LDGLQVEAEVVTNGTSCSESTTQNDGDCRTRDTRVSTNGSDPEVAAS 178

```

```
146 G-----LLENNEGVYEDSGPGLSCFMEEAPYT-----DSTG 179
      |      |      |      |      |      |      |      |      |      |
179 GENKANGNNSPSLNGG--FKPSRPPRP-----SRPPPTPRPASVNGSPSTNSDSDG 231
      |      |      |      |      |      |      |      |      |      |
180 AA-----AGGNCRFVESPSQ-----DORLQAR 203
      |      |      |      |      |      |      |      |      |      |
232 SSTGSLPPTNTVNTSTSGATGIIILITSGSGPRPLNTVSOALPQWEGRV-----287
      |      |      |      |      |      |      |      |      |      |
204 LRNPDPVRLS-----QTPQNRPHGHSPELPEGYEQRTTVQGVYFLHTQTGVSTWH 255
      |      |      |      |      |      |      |      |      |      |
288 ----DQGRVYVVDHVEKRTTWDRE-----PLPGWERRVDNMGRIYVVDHFTRTTQ 338
      |      |      |      |      |      |      |      |      |      |
256 DRIP-----RDJ-----NSVNDGLGPPGWEVR 281
      |      |      |      |      |      |      |      |      |      |
339 RPTLESVRNYEQWLORSLOQAMQFQNRFYNGQDLFATSONKEFDPGLPFGWEKR 338
      |      |      |      |      |      |      |      |      |      |
282 STVSGRIYFVDHNRRTTQTPRLHHMHQQLKEPSQIPL-----324
      |      |      |      |      |      |      |      |      |      |
399 TDSNGRVYVFNHNRITQWEDR-----SQQLNE--KPLPEGWEMRFTVDGIPYFVDH 450
      |      |      |      |      |      |      |      |      |      |
325 -----PSEGSLEDELPQRYERDLVQKLVLK---HELSQLQOAGHCRIEVSRE 372
      |      |      |      |      |      |      |      |      |      |
451 NRATTYIDERTGKSALDNGPOIAYRDFKAKVQYFRFWCQQLAMPQ---HIKITVTRK 506
      |      |      |      |      |      |      |      |      |      |
373 EIPESYQIOMKRPDKLKRIMVFRGEEGLDYGVAREWLYLLCHEMLNPPYGLFOYS 432
      |      |      |      |      |      |      |      |      |      |
507 TLPEDSFQIMSFSPDLRRRLVWVFPSEGLDYGVAREWFFLLSHEVLNPMYCLPEYA 566
      |      |      |      |      |      |      |      |      |      |
433 TDNIYMLQNPSSINPDHLSVTFHVGIMGLAVPHGYINGFTVPYKOLLGKPIOLS 492
      |      |      |      |      |      |      |      |      |      |
567 GKDNICYLOINPASYINPDHLYFRFGRFIAMALPHGKFIDTGFSLPYKRLNKPVLGK 626
      |      |      |      |      |      |      |      |      |      |
493 DLBSVDPELHKSILWLENDITPV-LDHTFCVEHNAFGRILQHELKPNGRVNPVTENKX 551
      |      |      |      |      |      |      |      |      |      |
627 DLESIDPEFINSILWKENNIECGLEMYFSVDKEILGEIKSHDLKPNGGNLTVEENKE 686
      |      |      |      |      |      |      |      |      |      |
552 EYVRLVYVNRFRGIGIAQFLALQGFNELIPOHLLKPFDPQKELELIIGGLDKIDLNDWKS 611
      |      |      |      |      |      |      |      |      |      |
687 EYIRVMAEWRLSRGVEEQAFPEGFNEILPOYLQYFDAKELEVLGCGMEIDLNDWQR 746
      |      |      |      |      |      |      |      |      |      |
612 NTELKCVADSNIVRFQAVTTFDEERARLLQFVGTSTVPLQGEKALOGSGAGR 671
      |      |      |      |      |      |      |      |      |      |
747 HAIYRHYTTSQIMWFQVFEKIDEXRMRLQLQVGTGCRPLVGGFADLMGNS---GPQ 803
      |      |      |      |      |      |      |      |      |      |
672 LFTIHLIDANTNLPKATCFNRIIDIPVESYKELYKLLTAVBETCGFAVE 723
      |      |      |      |      |      |      |      |      |      |
804 KFCIEKV-GKENWLPESHTCFNRLDLPYKSYEQLEKLLFAIETETEGFGOE 854
      |      |      |      |      |      |      |      |      |      |
Q9BY75 PRELIMINARY; PRT; 862 AA.
Q9BY75;
01-JUN-2001 (T-EMBLrel. 17, Created)
01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
Ubiquitin protein ligase Itch.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
Miyazaki K., Okamoto Y., Sakamoto M., Nakagawara A.;
"Homo sapiens mRNA for ubiquitin protein ligase Itch, complete cds."
Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
[2]
SEQUENCE FROM N.A.
MEDLINE=20549573; PubMed=10840313;
Qiu L., Joazeiro C., Fang N., Wang H.Y., Elly C., Altman Y., Fang D.,
Hunter T., Liu Y.C.;
```

```
RT Recognition and ubiquitination of Notch by Itch, a hect-type E3
RL ubiquitin ligase."
RN J. Biol. Chem. 275:35734-35737(2000).
RP SEQUENCE FROM N.A.
RX MEDLINE=98122574; PubMed=9462742;
RA Perry W.L., Hustad C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,
Copeland N.G.;
RT "The Itchy locus encodes a novel ubiquitin protein ligase that is
RL disrupted in a18H mice."
RN Nat. Genet. 18:143-146(1998).
RP SEQUENCE FROM N.A.
RX MEDLINE=20501262; PubMed=11046148;
RA Winberg G., Matskova L., Chen F., Plant P., Rotin D., Gish G.,
Ingham K., Erberig I., Fawson T.;
RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
RL protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases."
RN Mol. Cell. Biol. 20:8526-8535(2000).
RP SEQUENCE FROM N.A.
RX MEDLINE=98313405; PubMed=9647693;
RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;
RT "Atrophin-1, the DRPLA gene product, interacts with two families of WW
RL domain-containing proteins."
RN Mol. Cell. Neurosci. 11:149-160(1998).
RP SEQUENCE FROM N.A.
RX MEDLINE=21219930; PubMed=11318614;
RA Chen X., Wen S.C., Fukuda M.N., Gavva N.R., Hsu D.W., Akama T.O.,
Yang-Peng T.L., Shen C.K.J.;
RT "Human ITCH is a Co-Regulator of the Hematopoietic Transcription
RL Factor NF-E2."
RN Genomics 73:238-241(2001).
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AB056663; BAB39389.1; -.
DR HSSP; Q13526; 1PIN.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW_domain.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
KW Ligase.
SQ SEQUENCE 862 AA; 98675 MW; A3D960E7F4DBF9D3 CRC64;
Query Match 33.5%; Score 1301; DB 4; Length 862;
Best Local Similarity 34.7%; Pred. No. 4.8e-98;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;
QY 6 KIRLTVLCAK-NLAKDFFRLDPDPFAKIVVDGSGQCHSTDTVNTLDPKWNQHYDLYVGK 64
      |      |      |      |      |      |      |      |      |      |
DB 19 QLQITVISAKLKENKNWFG-PSYVEVTD--GQSKTEKCNNTNSPKWQPLTIVTP 75
      |      |      |      |      |      |      |      |      |      |
QY 65 TDSITISVWNKKTHKK--QGAGFLGCVRLLSNAIRLKD-----GYQR-----L 108
      |      |      |      |      |      |      |      |      |      |
DB 76 VSKLHFRVWSHQTLKSDVLLGTAALDIYETIKSNMKLEEVVTVLQGGDKPTETIGL 135
      |      |      |      |      |      |      |      |      |      |
QY 109 DLCKLNPSDITDAVRGOIVWSLOT-----RDRIGTGS-----VWDC 144
      |      |      |      |      |      |      |      |      |      |
DB 136 SIC-----LDGLQLESEVVVTNGETTCSESASQNDGSRKDETVSTNGSDPDADAGEN 191
```



```

640 RARLLQFVTGSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTWNLPLKAHTCFNRIDIPF 699
824 RMALLQFVTGTCRLPVGGFADLMGNS--GPKFCIEKV-GKENWLPRSHTCFNRULDLP 879
700 YESVEKLYEKLITAVEETCGRAVE 723
880 KYSEYQLKEKLIFAIETEGFQE 903

RESULT 10
INSA7
  Q8NSA7 PRELIMINARY; PRT; 911 AA.
  Q8NSA7:
    01-OCT-2002 (T-EMBLrel. 22, Created)
    01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
    01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
    Similar to neural cell expressed, developmentally down-regulated
    4-like.
    Homo sapiens (Human).
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
    NCBI_Taxid=9606;
  [1]
    SEQUENCE FROM N.A.
    TISSUE=Melanoma;
    Strausberg R.;
    Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
    -I- SIMILARITY: CONTAINS 1 C2 DOMAIN.
    EMBL; BC032597; AH32597.1; -.
    InterPro; IPR000008; C2.
    InterPro; IPR000569; HECT domain.
    InterPro; IPR002349; WW.
    InterPro; IPR001202; WW_Rsp5_WWP.
    Pfam; PF00168; C2; 1.
    Pfam; PF00632; HECT; 1.
    Pfam; PF00397; WW; 3.
    PRINTS; PR00360; C2DOMAIN.
    PRINTS; PR00403; WWDOMAIN.
    SMART; SM00239; C2; 1.
    SMART; SM00119; HECTC; 1.
    SMART; SM00456; WW; 3.
    PROSITE; PS00499; C2 DOMAIN_1; 1.
    PROSITE; PS00004; C2 DOMAIN_2; 1.
    PROSITE; PS00237; HECT; 1.
    PROSITE; PS01159; WW DOMAIN_1; 3.
    PROSITE; PS00020; WW DOMAIN_2; 3.
    SEQUENCE 911 AA; I04921 NW; C504AED677AA506 CRC64;

Query March
Best Local Similarity 32.7%; Score 1269; DB 4; Length 911;
Matches 310; Conservative 125; Mismatches 246; Indels 248; Gaps 24

2 GSSIKIRLTVLCAKNLAKDFFRLDPPAKI---VVDGSGQCH--STDVKNLTLDPKWQ 56
17 GESRILEVKKVSGIDLAKKIDIGASDPYVKLSLYADENRELALVQFTTKTLNFKWNE 76
57 HYDLYVKTD-STISVWNKKHKHKKGAGFGVGLLSNAI----- 97
77 EFYFRVNPNSHRLLFVFDENLRITDD---FLGQVDVPLSLPDEPTMERPYTFKDFLL 133
98 -----SRLKDTQGYRLDCKLNLPDSDTDAVGQIVSLQTEDRIGTGSVVDCRG----- 146
134 RPSRHSRKVK--GFLRLKWAYPKNG-----GQDEENSQDDMEHGEVVDNSASQH 186
147 -----LLENEGTVY----- 155
187 QEELPPPLPPFGWEKVDNLGRITYVYVHNHRTTQWHRPSLMDVSSSDNNIRQINQEAHH 246
156 -----EDSGP-----GRPLSCFWEAPAYTDSGTGA-----AGGNCRFVES 192
247 RFRSRRHISEDLEPPSGGVDVPFWETIIEEVNVIAGDSLGLALPPPPASPGSRSTPQE 306
193 PSQD-----ORLQAQLRN----- 206

```



```
157 PE-----PLPPGKERRVDNMGRYIYVDHFTRTTWTQRTLESVRYNQWLQSRQLQAGM 211
261 -----RDL-----NSVNCDELGLPPLPGWEVRSVTSGRYIYVDHNNRTTQFTDRL 305
212 QQPNORFIYGNQDLFATSQKSFDFGLPPLPGWEKRTDSNGRVYFVNHNTITOWEDPR- 270
306 HHIMNECCOLKPSQPLPL-----PSEGSLEDEELPAQR 339
271 -----SQQLNE--KPLPGWEVRFVTDGIPYFVDHNNRTTYIDPRTGKSLDNGFQIA 323
340 YERDLVQKLVLR---HELSLQOQAGHCRIEVSREBEIFESYRQIMKMRPKDLKRLMW 396
324 YVRDFKAKVQYFRFWCQQLAMPQ---HIKITVTRKTLFEDSFQOIMSFSPQDLRRRLMW 379
397 KPRGEGLDYGGVAREWYLLCHEMLNPYGLFOYSTDNIYMLQINPDSINPDHLSYFH 456
380 IFPEEGLDYGGVAREWFFLLSHEVLNPMYCLFAYAGKNTCLQINPASYINPDHLKYFR 439
457 FVGRIMGLAVFHGHYINGGFTVPFYKQLGKPIQLSDLESVDPELHKSLLVWLENDITPV 516
440 FIGRFIAMLFGHFKFIDTGSPLFYKRLINKPVGLKLESIDPEFYNSLIWKENNIEEC 499
517 -LDHTFCVEHNAFRILQHELKPNGRNVPVTENKKEYVRLYVNWRFMRGIEAQFLALQK 575
500 DLEMYSFVDKEILGEIKSHDLKPNNGNLTVTENKKEYIRVMAEWRLSRGVEBQTQAFPE 559
576 GFNELIPOHLLKPFQDKLELIGGLDKIDLNDWKSNTRLKHCVADSNIVRWFQAVET 635
560 GFNEILPQOYLQYFDAKELEVLGCGQBEIDLNDWQRHAIYRHARTSKQIMWFQVKEI 619
636 DEERRARLLQFVTGSTRVLPQFKALQSGTGAAGPRLLTIHLIDANTDNLPKAHTCFNRI 695
620 DNEKRWLLQFVTGTRCLRPVGGFADLNGSN---GPOKFCIEKV-GKENWLPRSHTCFNRL 675
696 DIPPYSEYKLEKLLTAVEETCGFAVE 723
676 DLPPYKSYEQLEKLLFAETEETGFGQE 703

O43584 PRELIMINARY; PRT; 739 AA.
3584
SEQUENCE FROM N.A.
MEDLINE=98313405; PubMed=9647693;
Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.:
"Atrophin-1, the DRPLA gene product, interacts with two families of WW
domain-containing proteins.";
Mol. Cell. Neurosci. 11:149-160 (1998).
EMBL; AF038564; AAC04845.1; -;
HSPB; Q13526; IPIN.
InterPro; IPR000569; HECT_domain.
InterPro; IPR002349; WW.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 4.
PRINTS; PR00403; WWDOMAIN.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 3.
PROSITE; PS50237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 4.
PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
FT NON_TER 1 1
SQ SEQUENCE 739 AA; 85057 MW; 7F1B4B8C105951B CRC64;
Query Match 32.1%; Score 1245.5; DB 4; Length 739;
Best Local Similarity 39.8%; Pred. No. 1.4e-93;
Matches 274; Conservative 99; Mismatches 180; Indels 135; Caps 18;
QY 114 NPSDTDAVRGQIVVSL-----QTRDRIG-TGGSVVDCRGLLENEGTVYEDSGPRPLSCF 167
DB 109 SFSATSESDGSTGSLSPPTNTNTSEGATSGLIPL-----TISGSGP-RPLNPV 159
QY 168 MEEPAPYTDSTGAAAGGNCRFVESPQDQRLQAOQLRNPDPVRGSL-----QTPQNR 219
DB 160 TQAPLP-----PCWEQRV-----DQGRVYVVDHVEKRTTWR 192
QY 220 PHGQSPPLPGYQRTTVQGVYFHTQTGVSTWHDPRIP----- 260
DB 193 PE-----PLPPGWEVRRVDNMGRYIYVDHFTRTTWTQRTLESVRYNQWLQSRQLQAGM 247
QY 261 -----RDL-----NSVNCDELGLPPLPGWEVRSVTSGRYIYVDHNNRTTQFTDRL 305
DB 248 QQPNORFIYGNQDLFATSQKSFDFGLPPLPGWEKRTDSNGRVYFVNHNTITOWEDPR- 306
QY 306 HHIMNQOLKEPSQPL-----PSEGSLEDEELPAQR 339
DB 307 -----SQQLNE--KPLPGWEVRFVTDGIPYFVDHNNRTTYIDPRTGKSLDNGFQIA 359
QY 340 YERDLVQKLVLR---HELSLQOQAGHCRIEVSREBEIFESYRQIMKMRPKDLKRLMW 396
DB 360 YVRDFKAKVQYFRFWCQQLAMPQ---HIKITVTRKTLFEDSFQOIMSFSPQDLRRRLMW 415
QY 397 KPRGEGLDYGGVAREWYLLCHEMLNPYGLFOYSTDNIYMLQINPDSINPDHLSYFH 456
DB 416 IFPEEGLDYGGVAREWFFLLSHEVLNPMYCLFAYAGKNTCLQINPASYINPDHLKYFR 475
QY 457 FVGRIMGLAVFHGHYINGGFTVPFYKQLGKPIQLSDLESVDPELHKSLLVWLENDITPV 516
DB 476 FIGRFIAMLFGHFKFIDTGSPLFYKRLINKPVGLKLESIDPEFYNSLIWKENNIEEC 535
QY 517 -LDHTFCVEHNAFRILQHELKPNGRNVPVTENKKEYVRLYVNWRFMRGIEAQFLALQK 575
DB 536 DLEMYSFVDKEILGEIKSHDLKPNNGNLTVTENKKEYIRVMAEWRLSRGVEBQTQAFPE 595
QY 576 GFNELIPOHLLKPFQDKLELIGGLDKIDLNDWKSNTRLKHCVADSNIVRWFQAVET 635
DB 596 GFNEILPQOYLQYFDAKELEVLGCGQBEIDLNDWQRHAIYRHARTSKQIMWFQVKEI 655
QY 636 DEERRARLLQFVTGSTRVLPQFKALQSGTGAAGPRLLTIHLIDANTDNLPKAHTCFNRI 695
DB 656 DNEKRWLLQFVTGTRCLRPVGGFADLNGSN---GPOKFCIEKV-GKENWLPRSHTCFNRL 711
QY 696 DIPPYSEYKLEKLLTAVEETCGFAVE 723
DB 712 DLPPYKSYEQLEKLLFAETEETGFGQE 739

RESULT 14
Q9NT88
ID Q9NT88 PRELIMINARY; PRT; 820 AA.
AC Q9NT88;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKF2P434P2422
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
```

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

EMBL; ALI37469; CAB70754.1; -

HSP; Q13526; 1PIN.

InterPro; IPR000008; C2.

InterPro; IPR000569; HECT_domain.

InterPro; IPR002349; WW.

InterPro; IPR001202; WW_Rsp5_WWP.

Pfam; PF00168; C2; 1.

Pfam; PF00632; HECT; 1.

Pfam; PF00397; WW; 3.

PRINTS; PR00403; WWDOMAIN.

SMART; SM00239; C2; 1.

SMART; SM00119; HECTC; 1.

SMART; SM00456; WW; 3.

PROSITE; PS00004; C2_DOMAIN_2; 1.

PROSITE; PS0237; HECT; 1.

PROSITE; PS01159; WW_DOMAIN_1; 3.

PROSITE; PS00020; WW_DOMAIN_2; 3.

Hypothetical protein.

NON TER

SEQUENCE 820 AA; 95283 MW; 0FDB34B29B3F4123 CRC64;

Query Match

Best Local Similarity 32.0%; Score 1241.5; DB 4; Length 820;

Matches 294; Conservative 116; Mismatches 230; Indels 203; Gaps 21;

43 TDVTKNTLPKNOHVDLVGKTD-SITISVWNKKIHKQAGFLGCVRLLSNAI----- 97

12 TKTIKTLAPKNEEYFVFNSENHLLFEVDENLRTRDD---FLGQVDVPLSHLTED 68

98 -----SRKGTGYQRLDCKLNPSDSTDVAVGQIVSVLQTRDRIGT 137

69 PTMERPYTFKDFLLRPRSHKSRVK-GFLRLKWAYMPKNG-----GQDEENSDDRDDMEH 121

138 GGSVVDRCG-----LLENEGTVY----- 155

122 GWEVDSNDASQHQELPPPLPPQWSEKVDNLGRTYYVNNHNTTQWHRPSLMDVSSE 181

156 -----EDSGP-----GRPLSCFMEEPAPTYDSTGAA-- 181

182 SDNNIRQINQEAHRRFRRRHISELEPESEGGDVPEPWETISEEVNIAGDSLGLALP 241

182 ---AGGNCRFVESPDQ-----ORLQAQRLN----- 206

242 PPASPSPSRTPQELSEELSRRLQITPDNGSQFSLIQREPSLSRSCSVTDVAEGQH 301

207 -----PDVRGSLQTPQNRPHGHQSPE-----LPEGYQRITVQGVYFL 245

302 LPPGAKDSPVRRAVKDTLSNPQSPQSPYNSPKQKHVQTSFLPPGWMRIAPNGRPFI 361

246 HTQTGVSTWHDPRIP-----RDLSVNCDELGLPLPGWEVRSVSGRIYFDVHNNRTQF 300

362 DHNTKTKTTWEDPLKFPVHMRSKTSLNPNLGLPLPGWEERIHLDGRTFYFDHNSKITQW 421

301 TDPLRLHMHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLVQKLVRHESLQOP 360

422 EDPLQN-----PAITG-----PAVYSEFPKQDYFKKLKPAD 458

361 QAGHCRIEVSRBIFESYRQIMM-RPKDLKRLMKVPRGEGLDYGVAREWLYLICH 419

459 IPNRFEKMLHNNIFESYRRIRMSVRPDVLRKARLWIEPESEKGLDYGAVAREWPFLLSK 518

420 EMLNPFYGLFOYSDNIMLQINPDSSI-NPDHLSYFHFVGRIMGLAVFHGYINGFTV 478

519 EMFNPFYGLFEYATNTYLQINPNGLCNEDHLSYFTIGVAGLAVFHGKLLDGFIR 578

479 PFYKQLGKPIQLSDLESVDPELHKLWILENDITPVLDTHTFCVHEHNAFGRILQHELP 538

579 PFYKMWLGKQITLNDMSVDSEYNSLKWLENDPTE-LDLMEFCIDEENFGQTYQVDLKP 637

539 NGRVNPVTENKEVYRLVYNTWFMRGIEAQFLALQKGFNELIPQHLKPPDKLELII 598

Db 638 NGSEIWTNENKREYIDLVIQWRFVNRVQKMAFLEGTELLPIDLIKIFDNELELLM 697

QY 599 GLDKIDLDNWKSNRLKHKVADSN-IVRFWQAVETFEERPARLLQFVGTGSTRVPLOG 657

Db 698 CGLGDVDVNDWRQHSIYKNGYCPNHPVIOFWKAVLMDAEKRIQLQFVGTGSTRVPMNG 757

QY 658 FKALGGSTGAAGPRLFTIHLIDANTDLNPKAHTCFNRIDIPPYSEYKLYEKLITAVEET 717

Db 758 FAELYGSN---GPQLFTIEQW-GSPEKLPRAHYCFNRLDLPPIETFEDELEKLLMAVENA 813

QY 718 QGF 720

Db 814 QGF 816

RESULT 15

Q9HOMO PRELIMINARY; PRT; 922 AA.

AC Q9HOMO; DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)

DE Hypothetical protein (WW domain-containing protein 1).

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1] TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=21154917; PubMed=11230166;

RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,

RA Ansorge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,

RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,

RA Mewes H.W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,

RA Wambutt R., Korn B., Klein M., Poustka A.;

RT Towards a Catalog of Human Genes and Proteins: Sequencing and

RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.;

RL Genome Res. 11:422-435(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Flaszka M., Aslam H., Poulson R., Hanby A., Gorman P., Roylance R.,

RA Nolan T., Brady G., Canfield A., Baron M.;

RT "WPI", a human homolog of Drosophila suppressor of dextex.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

DR EMBL; ALI36739; CAB66673.1; -

DR EMBL; AY043361; AAK94668.1; -

DR EMBL; BC036065; AAK36065.1; -

DR HSP; Q13526; 1PIN.

DR InterPro; IPR000008; C2.

DR InterPro; IPR000569; HECT_domain.

DR InterPro; IPR002349; WW.

DR InterPro; IPR001202; WW_Rsp5_WWP.

DR Pfam; PF00168; C2; 1.

DR Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 4.

DR PRINTS; PR00403; WWDOMAIN.

DR SMART; SM00239; C2; 1.

DR SMART; SM00119; HECTC; 1.

DR SMART; SM00456; WW; 4.

DR PROSITE; PS00004; C2_DOMAIN_2; 1.

DR PROSITE; PS0237; HECT; 1.

DR PROSITE; PS01159; WW_DOMAIN_1; 4.

DR PROSITE; PS00020; WW_DOMAIN_2; 2.

DR Hypothetical protein.

KW SEQUENCE 922 AA; 105201 MW; 35B6E1C03A3147DA CRC64;

arch completed: February 20, 2004, 15:29:53
b time : 40.9137 secs

N-PSDB; AAF24953.

Novel isolated Smurf protein useful for inhibiting bone morphogenic protein or tumor growth factor-beta activation pathway, for treating cancer and to block osteogenesis, hair growth, tooth formation - Claim 10; Fig 12; 107pp; English.

The present sequence represents a human Smurf2 polypeptide. The specification also describes a Smurf1 polypeptide. Smurf polypeptides are negative regulators of Smad signal transduction, and antagonists of bone morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta) signalling pathway. Expression of Smurf1 in a cell is useful for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf polypeptides are useful for blocking chondrogenesis, osteogenesis, blood differentiation, cartilage formation, neural tube patterning, retinal development, heart induction and morphogenesis, hair growth, tooth formation, gamete formation and a wide variety of tissue and organ formation processes, and hinder the regeneration, growth, maintenance, etc., of bone and other tissues that are dependent on the BMP pathway. The polypeptide is useful for screening for various drugs and/or antibodies that can either enhance the BMP pathway, or inhibit it.

Sequence 748 AA;

Query Match 100.0%; Score 4038; DB 22; Length 748;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSNPGRRNGPVKRLTLVLCANLVKDKDFRLDPFAKVVVDGSGQCHSTDTVKNLTDPK 60
1 MSNPGRRNGPVKRLTLVLCANLVKDKDFRLDPFAKVVVDGSGQCHSTDTVKNLTDPK 60
61 WNOHYDLIYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDCKL 120
61 WNOHYDLIYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDCKL 120
121 GPNDNDTVRGQIVVLSQSDRIGTGGQVVDGSRFLPDNDLPDGEERRTASGRIOVLNHT 180
121 GPNDNDTVRGQIVVLSQSDRIGTGGQVVDGSRFLPDNDLPDGEERRTASGRIOVLNHT 180
181 RTTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGQSSDPRLAERRVSQRHN 240
181 RTTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGQSSDPRLAERRVSQRHN 240
241 YMSRTHLTPDPLPEGVEORTTQGGQVYFLHTQTGVSTVHWDPRVPRDLNINCELGFLP 300
241 YMSRTHLTPDPLPEGVEORTTQGGQVYFLHTQTGVSTVHWDPRVPRDLNINCELGFLP 300
301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRQNLKXQDQVSLCPDDT 360
301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRQNLKXQDQVSLCPDDT 360
361 ECLTVPRYKDLVQKLILQBELSQOQPAQCHRIEVSREEIFEESYQVWMPKDLWK 420
361 ECLTVPRYKDLVQKLILQBELSQOQPAQCHRIEVSREEIFEESYQVWMPKDLWK 420
421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNYPYGLFOYSRDDIVTLQINPDVANPEHL 480
421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNYPYGLFOYSRDDIVTLQINPDVANPEHL 480
481 SYHFVGRIMGMAVFHGHYIDGGFTLFFYKQLLGKSIITLDDMELVDPDLNLSVWILEND 540
481 SYHFVGRIMGMAVFHGHYIDGGFTLFFYKQLLGKSIITLDDMELVDPDLNLSVWILEND 540
541 ITGVLDHTFCVENNAVGEIITQHELKNGKSIIPVNEENKKEYVRLVYNRFLRGTEAQFLA 600
541 ITGVLDHTFCVENNAVGEIITQHELKNGKSIIPVNEENKKEYVRLVYNRFLRGTEAQFLA 600
601 LQKGFNEVIFQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKMFKAV 660
601 LQKGFNEVIFQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKMFKAV 660

QY 661 EFDERRARLLQFVTGSSRVPLQGFALQAGAPRLFTTHQIDACTNNLPKHAHTCFNRI 720
Db 661 EFDERRARLLQFVTGSSRVPLQGFALQAGAPRLFTTHQIDACTNNLPKHAHTCFNRI 720
QY 721 DIPPYESYEKLYEKLTAIETTCGFAVE 748
Db 721 DIPPYESYEKLYEKLTAIETTCGFAVE 748

RESULT 2

AAW13384
ID AAW13384 standard; Protein; 748 AA.
XX AAW13384;
AC AAW13384;
DT 10-JUL-1997 (first entry)
XX Human protein ubiquitin ligase publ.
DE Protein ubiquitin ligase; publ; cdc25 phosphatase; CDK kinase;
XX p53; cell cycle; transgenic animal.
KW p53; cell cycle; transgenic animal.
XX Homo sapiens.
OS Homo sapiens.
XX WO9712962-A1.
PN 10-APR-1997.
XX 04-OCT-1996; 96WO-US15930.
XX 04-OCT-1995; 95US-0539205.
XX (COLD-) COLD SPRING HARBOR LAB.
XX Beach D, Caligiuri M, Nefsky B;
PI WPI; 1997-226206/20.
DR N-PSDB; AAT47040.
XX Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved
PT in cdc25 phosphatase and p53 ubiquitination, and regulate cell
PT growth and proliferation
XX Claim 1; Page 74-77; 108pp; English.
XX Human protein ubiquitin ligases publ (AAW13384), pub2 (AAW13385) and
CC pub3 (AAW1386) are homologues of fission yeast pub1 (AAW13387) and
CC were identified from cDNA clones (AAT47040-42) obtcd. e.g. from a
CC keratinocyte cDNA library. Pub polypeptides can be produced in
CC transfected host cells. They can control the steady state level of
CC cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2)
CC dephosphorylation and the steady state level of p53 (controlling
CC the degree of cell cycle regulation of p53). They can regulate
CC cell or tissue differentiation, or cell growth or proliferation by
CC affecting other proteins, can be a specific (ant)agonist of wild-
CC type protein function and may be used as immunogens to elicit a
CC specific immune residue.
XX Sequence 748 AA;
SQ

Query Match 99.8%; Score 4031; DB 18; Length 748;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSNPGRRNGPVKRLTLVLCANLVKDKDFRLDPFAKVVVDGSGQCHSTDTVKNLTDPK 60
Db 1 MSNPGRRNGPVKRLTLVLCANLVKDKDFRLDPFAKVVVDGSGQCHSTDTVKNLTDPK 60
QY 61 WNOHYDLIYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDCKL 120
Db 61 WNOHYDLIYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDCKL 120
QY 121 GPNDNDTVRGQIVVLSQSDRIGTGGQVVDGSRFLPDNDLPDGEERRTASGRIOVLNHT 180

Query Match	99.7%;	Score 4027;	DB 22;	Length 804;	
Best Local Similarity	99.7%;	Pred. No. 0;	Mismatches 2;	Indels 0;	Gaps 0;
Matches 746;	Conservative 0;				
1	MSNPGRRRGPKVRLTLVLCANLVKQDFRLLDPFAKVVVDGSGQCHSTDTVXNTLDPK	60			
57	MSNPGRRRGPKVRLTLVLCANLVKQDFRLLDPFAKVVVDGSGQCHSTDTVXNTLDPK	116			
61	WNOHYDLYTGKSDSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL	120			
117	WNOHYDLYTGKSDSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL	176			
121	GPNDNDTVRGQIVVSLQSRDRIGTGGQVVDCCSRLFDNDLPDQWEERRTASGRIOYLNHIT	180			
177	GPNDNDTVRGQIVVSLQSRDRIGTGGQVVDCCSRLFDNDLPDQWEERRTASGRIOYLNHIT	236			
181	RTTQWERPTPASEYSSPGRPLSCFVDENTPISGTNGATCGQSSDPRLAERRVRSQRHN	240			
237	RTTQWERPTPASEYSSPGRPLSCFVDENTPISGTNGATCGQSSDPRLAERRVRSQRHN	296			
241	YMSRTHLTPPDLPEGYEQRTTQOGQVYFLHTQTGVSTWHPDPRVPRDLSNINCELGPLP	300			
297	YMSRTHLTPPDLPEGYEQRTTQOGQVYFLHTQTGVSTWHPDPRVPRDLSNINCELGPLP	356			
301	PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVVSCLPDDT	360			
357	PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVVSCLPDDT	416			
361	ECLTVPRYKEDLVQKLKILROELSQQOQAGHCRIEVSREEIFEESYQVWMPKDLWK	420			
417	ECLTVPRYKEDLVQKLKILROELSQQOQAGHCRIEVSREEIFEESYQVWMPKDLWK	476			
421	RLMKFRGEGLDYGVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL	480			
477	RLMKFRGEGLDYGVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL	536			
481	SYPHFVGRIMGMAVFGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND	540			
537	SYPHFVGRIMGMAVFGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND	596			
541	ITGVLDHTFCVHNAYGEIIOHELKPNKGSIPVNEENKEYVRLYVNNWRLRGIEAFLA	600			
597	ITGVLDHTFCVHNAYGEIIOHELKPNKGSIPVNEENKEYVRLYVNNWRLRGIEAFLA	656			
601	LQGFNEVIPQHLTKTFDEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWPKAV	660			
657	LQGFNEVIPQHLTKTFDEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWPKAV	716			
661	EFFDEERRARLLQFVTGSSRVLPQGFALQGAAGPRLFTIHOIDACTNNLPAKHTCFNRI	720			
717	EFFDEERRARLLQFVTGSSRVLPQGFALQGAAGPRLFTIHOIDACTNNLPAKHTCFNRI	776			
721	DIPYBSYKLYEKLJTAIETCGFAVE	748			
777	DIPYBSYKLYEKLJTAIETCGFAVE	804			
Query Match	74.3%;	Score 3001;	DB 24;	Length 722;	
Best Local Similarity	74.6%;	Pred. No. 3.8e-274;			
Matches 565;	Conservative 63;	Mismatches 71;	Indels 58;	Gaps 8;	
QY	12 VKLRITVLCAKNLVKQDFRLLDPFAKVVVDGSGQCHSTDTVXNTLDPKWNQHYDLYGK	71			
Db	4 IKIRLITVLCAKNLVKQDFRLLDPFAKVVVDGSGQCHSTDTVXNTLDPKWNQHYDLYGK	63			
QY	72 SPSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	131			
Db	64 TDSITISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	123			
QY	132 IVVSLQSRDRIGTGGQVVDCCSRLFDNDLPDQWEERRTASGRIOYLNHITTTQWERPTP	191			
Db	124 IVVSLQSRDRIGTGGQVVDCCSRLFDNDLPDQWEERRTASGRIOYLNHITTTQWERPTP	154			
QY	192 ASEYSSPGRPLSCFVDENTPISGTNGATCG-----QSSDPRLAERRVRSQRHN	242			
Db	155 ---EDSGPGRPLSCFMEEPAPYDSTGAAAGGNCGRFVESFSQORLOAQRLNPDVRGSL	212			
QY	243 ----SRTHLTPPDLPEGYEQRTTQOGQVYFLHTQTGVSTWHPDPRVPRDLSNINCELG	298			
Db	213 QTPQNRKPHGHSPELPEGYEQRTTQOGQVYFLHTQTGVSTWHPDPRVPRDLSNINCELG	272			
QY	299 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVVS	354			
Db	273 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVVS	328			
QY	355 LCPDDTECLTVPRYKEDLVQKLKILROELSQQOQAGHCRIEVSREEIFEESYQVWMPK	414			
Db	329 L---EDEELPAQRYERDLVQKLKILROELSQQOQAGHCRIEVSREEIFEESYQVWMPK	385			
Query Match	74.3%;	Score 3001;	DB 24;	Length 722;	
Best Local Similarity	74.6%;	Pred. No. 3.8e-274;			
Matches 565;	Conservative 63;	Mismatches 71;	Indels 58;	Gaps 8;	
QY	12 VKLRITVLCAKNLVKQDFRLLDPFAKVVVDGSGQCHSTDTVXNTLDPKWNQHYDLYGK	71			
Db	4 IKIRLITVLCAKNLVKQDFRLLDPFAKVVVDGSGQCHSTDTVXNTLDPKWNQHYDLYGK	63			
QY	72 SPSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	131			
Db	64 TDSITISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	123			
QY	132 IVVSLQSRDRIGTGGQVVDCCSRLFDNDLPDQWEERRTASGRIOYLNHITTTQWERPTP	191			
Db	124 IVVSLQSRDRIGTGGQVVDCCSRLFDNDLPDQWEERRTASGRIOYLNHITTTQWERPTP	154			
QY	192 ASEYSSPGRPLSCFVDENTPISGTNGATCG-----QSSDPRLAERRVRSQRHN	242			
Db	155 ---EDSGPGRPLSCFMEEPAPYDSTGAAAGGNCGRFVESFSQORLOAQRLNPDVRGSL	212			
QY	243 ----SRTHLTPPDLPEGYEQRTTQOGQVYFLHTQTGVSTWHPDPRVPRDLSNINCELG	298			
Db	213 QTPQNRKPHGHSPELPEGYEQRTTQOGQVYFLHTQTGVSTWHPDPRVPRDLSNINCELG	272			
QY	299 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVVS	354			
Db	273 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVVS	328			
QY	355 LCPDDTECLTVPRYKEDLVQKLKILROELSQQOQAGHCRIEVSREEIFEESYQVWMPK	414			
Db	329 L---EDEELPAQRYERDLVQKLKILROELSQQOQAGHCRIEVSREEIFEESYQVWMPK	385			

ULT 4
32724

AAE32724 standard; Protein; 722 AA.

AAE32724;

24-MAR-2003 (first entry)

E3 ubiquitin ligase SMURF1 protein.

Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis; poliomyelitis; HIV; measles; protein therapy; E3 ubiquitin ligase; enzyme.

Unidentified.

415 PKDLKRLMIKPRGEGGLYGGVAREWMLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 474
 386 PKDLKRLMIKPRGEGGLYGGVAREWMLYLLSHEMLNPPYGLFOYSTONIMYQLINPDS 445
 475 VNPHELSYFHFVGRIMGMAVFFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPOLHNSLV 534
 446 INPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHNSLV 505
 535 WILENDITGVLDHTFCVEHNAVEIIOHELKPNKSIIPVNEENKKEYVRLVYVNRPLRGI 594
 506 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLVYVNRFRMGI 565
 595 EAQFLALQKGFNEVIPPQHLIKPFDKELELIICGLGKIDVNDKVNTRKHKCTPDSNIVK 654
 566 EAQFLALQKGFNELIPQHLIKPFDQKLELIIGGLDKIDLNDKSNTRKHKCVADSNIVR 625
 655 WFKVAVEFDEERRARLLQFVTGSSRVPLOQFKALO---GAAGPRLFTIHQIDACTNNLP 711
 626 WFKVAVEFDEERRARLLQFVTGSTRVPLQGFALQGGSTGAAGPRLFTIHLIDANTNLP 685
 712 KAHTCFNRIDIPPYESYEKLYEKLTAIEETCGFAVE 748
 686 KAHTCFNRIDIPPYESYEKLYEKLTAIEETCGFAVE 722

ULT 5
 31476
 AAB31476 standard; Protein; 723 AA.

AAB31476;

20-APR-2001 (first entry)

Amino acid sequence of a human Smurf1 polypeptide.

Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP; transforming growth factor-beta; human; TGF-beta; chondrogenesis; osteogenesis; blood differentiation; cartilage formation; hair growth; neural tube patterning; retinal development; heart induction; morphogenesis; tooth formation; gamete formation.

Homo sapiens.

WO2000077168-A2.

21-DEC-2000.

12-JUN-2000; 2000WO-US16250.

11-JUN-1999; 99US-0138969.

(UTNY) UNIV NEW YORK STATE RES FOUND.
 (HSCR-) HSC RES & DEV LP.

Thomsen GH, Wrana J;

WPI; 2001-071267/08.
 N-P8DB; AAF24852.

Novel isolated Smurf protein useful for inhibiting bone morphogenic protein or tumor growth factor-beta activation pathway, for treating cancer and to block osteogenesis, hair growth, tooth formation -

Claim 6; Fig 10; 107pp; English.

The present sequence represents a human Smurf1 polypeptide. The specification also describes a Smurf2 polypeptide. Smurf polypeptides are negative regulators of Smad signal transduction, and antagonists of bone morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta) signalling pathway. Expression of Smurf1 in a cell is useful for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf polypeptides are useful for blocking chondrogenesis, osteogenesis, blood differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth formation, gamete formation and a wide variety of tissue and organ CC formation processes, and hinder the regeneration, growth, maintenance, CC etc., of bone and other tissues that are dependent on the BMP pathway. CC The polypeptide is useful for screening for various drugs and/or CC antibodies that can either enhance the BMP pathway, or inhibit it.
 XX Sequence 723 AA;
 SQ
 Query Match 74.3%; Score 3001; DB 22; Length 723;
 Best Local Similarity 74.6%; Pred. No. 3.9e-274;
 Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;
 QY 12 VKLRITVLCARLNVKDFRLLPDPFAKVVVDGSGQCHSTDTVKNTLDPKKNQHYDLYIGK 71
 DB 5 IKIRITVLCARLNVKDFRLLPDPFAKVVVDGSGQCHSTDTVKNTLDPKKNQHYDLYIGK 64
 QY 72 SSVTISVNNHKKHKKQAGFLGCVRLLSNAIRLKDGTGYQRLDCKLGFNDNDTVRGQ 131
 DB 65 TDSITISVNNHKKHKKQAGFLGCVRLLSNAIRLKDGTGYQRLDCKLGFNDNDTVRGQ 124
 QY 132 IVVLSQSDRIGTGGVVDCSRFPNDLPDQWEERBTASGRIOVLNHIITTTQWERTRP 191
 DB 125 IVVLSQSDRIGTGGVVDCSRFPNDLPDQWEERBTASGRIOVLNHIITTTQWERTRP 155
 QY 192 ABEYSRPGRLSCFVDENTPISGTNGATCG-----QSSDPRLAERRVRSQRHRYM 242
 DB 156 --EDSGPRPLSCFMEEPAPVTDSTGAAGGNCRCFVESPDSQDQRLQAQLRNPDRGSL 213
 QY 243 ----SRTHLHPPDLPEGYEORTTCOGVYELHTQTGVSTWHDPRVPRDLSNCELGP 298
 DB 214 QTFQNRPHGHOSPPELPEGYEORTTCOGVYELHTQTGVSTWHDPRVPRDLSNCELGP 273
 QY 299 LPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQQLKXQQQVW----S 354
 DB 274 LPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQQLKXQQQVW----S 329
 QY 355 LCPDDETECLTVPRYKRLVQKILRQLSQQQOAGHCRIEVSREIPEESYRQVMKR 414
 DB 330 L---EDEELPAQRYERDLVQKILRQLSQQQOAGHCRIEVSREIPEESYRQVMKR 386
 QY 415 PKDLKRLMIKPRGEGGLYGGVAREWMLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 474
 DB 387 PKDLKRLMIKPRGEGGLYGGVAREWMLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 446
 QY 475 VNPHELSYFHFVGRIMGMAVFFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPOLHNSLV 534
 DB 447 INPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHNSLV 506
 QY 535 WILENDITGVLDHTFCVEHNAVEIIOHELKPNKSIIPVNEENKKEYVRLVYVNRPLRGI 594
 DB 507 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLVYVNRFRMGI 566
 QY 595 EAQFLALQKGFNEVIPPQHLIKPFDKELELIICGLGKIDVNDKVNTRKHKCTPDSNIVK 654
 DB 567 EAQFLALQKGFNELIPQHLIKPFDQKLELIIGGLDKIDLNDKSNTRKHKCVADSNIVR 626
 QY 655 WFKVAVEFDEERRARLLQFVTGSSRVPLOQFKALO---GAAGPRLFTIHQIDACTNNLP 711
 DB 627 WFKVAVEFDEERRARLLQFVTGSTRVPLQGFALQGGSTGAAGPRLFTIHLIDANTNLP 686
 QY 712 KAHTCFNRIDIPPYESYEKLYEKLTAIEETCGFAVE 748
 DB 687 KAHTCFNRIDIPPYESYEKLYEKLTAIEETCGFAVE 723

RESULT 6

AAM79861

ID AAM79861 standard; Protein; 722 AA.

XX AAM79861;

XX AAM79861; (first entry)

DT 06-NOV-2001

Human protein SEQ ID NO 3507.

Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation.

Homo sapiens.

WO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US04098.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

20-JUN-2000; 2000US-0598075.

19-JUL-2000; 2000US-0620325.

01-SEP-2000; 2000US-0654936.

15-SEP-2000; 2000US-0663561.

20-OCT-2000; 2000US-0693325.

30-NOV-2000; 2000US-0728422.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y; Zhao QA, Wang P, Wang J, Zhang J, Ren F, Chen R, Wang ZH; Xue AJ, Yang Y, Wejhrman T, Goodrich R;

WPI; 2001-476283/51.

N-PSDB; AAK52994.

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy -
Claim 20; Page 362-363; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

Sequence 722 AA;

ery Match 74.1%; Score 2992; DB 22; Length 722;
st Local Similarity 74.5%; Pred. No. 2.7e-273;
atches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;

12 VKLRITVLCANLKVKKDFRLLPDPFAKVVVDGSGGCHSTDTVANTLDPKMNQHYDLYIGK 71

4 IKRLIVLCANLAKDKFRLLPDPFAKVVVDGSGGCHSTDTVANTLDPKMNQHYDLYIGK 63

72 SDSVTISVNNKKIHKKGAGFLGCVRLLSNAINRLKDTGTQRLDCLKGPNNDNTRGQ 131

64 TDSITISVNNKKIHKKGAGFLGCVRLLSNAINRLKDTGTQRLDCLKGPNNDNTRGQ 123

132 IVVSLOSRRDRIGTGGGVDCSRLFDNDLDPGWEERRTAGRIQYLNHITRTQWERPTRP 191

124 IVVSLQTRDRIGTGGGVDCRGLLENE-----GVVY----- 154

192 ASEYS8FGRLPSCFVDNTPISGTNGATCG-----QSSDPLAERRVRSORHRYM 242

Db 155 --EDSGEGRPLSCFMEEPAPYTDSTGAAGAGGNCRFVESPSODQLQAOQLANPDVRSGL 212
Qy 243 ----SRTHLHTPDLPEGEQRTTQGGQVYFLHTQGVSTWHDPRVPRDLNSINCEELGP 298
Db 213 QTPQNRPHGHSDELPEGEQRTTQGGQVYFLHTQGVSTWHDPRIPDLASVNCDELGP 272
Qy 299 LPFGWEIRNTATGRVYFVDHNETTQFTDPRLSANLHLVLAERONOLKDDQQQQQVV---S 354
Db 273 LPPGWEVRSTVSGRIYFVDHNRRTQFTDPR-----LHIMNHQCCQKEPSQDLPLPSGS 328
Qy 355 LCPDDTECLTVPRYKRDVLQKILRQELSQQQPGAGHCRIEVSREEIFEESYRQVMKOR 414
Db 329 L---EDBELPAQYERDLVQKLKLRHLSLQPPQAGHCRIEVSREEIFEESYRQIMKOR 385
Qy 415 PKDLKRLMIKFRGEGGLDYGVARWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDS 474
Db 386 PKDLKRLMIKFRGEGGLDYGVARWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDS 445
Qy 475 VNPEHLSYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLV 534
Db 446 INPDLSYFHFVGRIMGLAVFHHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHKSIV 505
Qy 535 WILENDITGVLDHTFCVEHNAIGEIIQHELKPKNGKSI PVNEENKKEYVRLYNNWFLRGI 594
Db 506 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVPTTEENKKEYVRLYNNWFLRGI 565
Qy 595 BAQFLAQKGFNVIPOHLLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPDSNIVK 654
Db 566 BAQFLAQKGFNELIPOHLLKTFDEKELELIIGLDKIDLDKNSNTRLKHCVADSNIVR 625
Qy 655 WFKXAVEFFDEERRARLLQFVTGSSRVPLQGGFKALQ---GAAGPRLFTIHQIDACTNMLP 711
Db 626 WFMQAVETDEERRARLLQFVTGSTRVPLQGGFKALQSGTGAAGPRLFTIHLIDANTDNL 685
Qy 712 KAHTCFNRIDIPPVESYEKLYEKLTAIBETCGFAVE 748
Db 686 KAHTCFNRIDIPPVESYEKLYEKLTAIBETCGFAVE 722

RESULT 7

AAU87301

ID AAU87301 standard; Protein; 514 AA.

AC AAU87301;

DT 05-JUN-2002 (first entry)

DE Novel central nervous system protein #211.

XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.

OS Homo sapiens.

PN WO200155318-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01332.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

366 GSTSGNGSQAAPQASASNGWTQEDAAITTSPTTTSPPRHSSQSPPTPNISPPASVTPSA 425
232 -----RVSQRHNYMSRTHLHTPP----- 251
426 NGNVHSPNANSTPAGSGGSRSYTAATPGQSRQRSSRQGBESSRRSSRGRNGGTS 485
252 -----DLPEGVEQRTTQGGVFLHTQTGVSTWHDPR 283
486 GGGGGGSGGQRYASAAATAANQAARPFLLPPGYEMRTTQGGVYFIHTPTGVSTWHDPR 545
284 VPRDL--SNINCEELGFLPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRQ 341
546 IPRDPDTQHLTDAIGLPLSGWEQRTASGRVYFVDHNNRTTQTDPRLSGSLQIRRG 605
342 N-----QLKQQQQQVSL-----CPDTE 361
606 TVPPTSANAATPAPPSPATPAGAAVFPQATPASNATPTLTTPPHHRIVPDLFQ 665
362 CL-----TVPRYKRLVQKILRQELSQQQPQAGHCRIEVSREEIFESYRQVMQRPK 416
666 GLEGADILPKYRRDLVGKRALATELTQWQPQSGHCRLEVSREIFESYRLIMQWRAK 725
417 DLKELMLKFRGEGLDYGGVAREWLYLLEHMLNPPYGLFQYSRDDIYTLQINPDSAVN 476
726 DMKELMYKFRGEGLDYGGVAREWLYLLEHMLNPPYGLFQYSRDDIYTLQINPDSGVN 785
477 PEHLSYFHFVGRIMGMVAFHGHYIDGGFTLPPFYKQLLGKSIITLDMELVDPDLHNSLWMI 536
786 PDHLSYFHF-----LNLKPITLGDIEGVDPDLHRLTWK 819
537 LENDITGVLDHTFCVEHNAYGEIIQHELKNGKSIPIVNEENKKEYVRLYVNMPLRGIEA 596
820 LESNTSGIESTFVSNNFSGALVHSLKPGGASIPVTEENKKEYVRLYVNMPLRGIEQ 879
597 QFALQKGFNEVIPHLLKTFDEKELELIGLGKIDVNDKYNLKHCTPDNSNVKWF 656
880 QFALQKGFCELIHSLHLPFDERELVIGGISSIDVNDWRNTRLKHCTNEITQVLWF 939
657 MKAVEFFDEERARLLQVTSRVPVLPQFKALQ--GAAGPLFTIH-QIDACTNNLPK 712
940 QWVESYSEMRARLLQVTSRVPVLPQFKALQSGTGAAGPLFTIHLDADVPTQNLK 999
713 AHTCFNRIDIPYSEYKELKLTATEETCGFAVE 748
1000 AHTCFNRIDIPYETQYLLCDKLTQAVEETCGFAVE 1035

JUL 9
113385

AAW13385 standard; Protein; 766 AA.

AAW13385;

10-JUL-1997 (first entry)

Human protein ubiquitin ligase pub2.

Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase;
p53; cell cycle; transgenic animal.

Homo sapiens.

WO9712962-A1.

10-APR-1997.

04-OCT-1996; 96WO-US15930.

04-OCT-1995; 95US-0539205.

(COLD-) COLD SPRING HARBOR LAB.

Beach D, Caligiuri M, Nefsky B;

XX WPI; 1997-226206/20.
DR N-PSDB; AAT47041.
XX Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved
PT in cdc25 phosphatase and p53 ubiquitination, and regulate cell
PT growth and proliferation
XX Claim 1; Page 80-84; 108pp; English.
XX Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and
CC pub3 (AAW13386) are homologues of fission yeast pub1 (AAW13387) and
CC were identified from cDNA clones (AAT47040-42) obt'd. e.g. from a
CC keratinocyte cDNA library. Pub polypeptides can be produced in
CC transfect host cells. They can control the steady state level of
CC cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2),
CC dephosphorylation and the steady state level of p53 (controlling
CC the degree of cell cycle regulation of p53). They can regulate
CC cell or tissue differentiation, or cell growth or proliferation by
CC affecting other proteins, can be a specific (ant)agonist of wild-
CC type protein function and may be used as immunogens to elicit a
CC specific immune residue.
XX Sequence 766 AA;
SQ Query Match 38.3%; Score 1545; DB 18; Length 766;
Best Local Similarity 40.8%; Pred. No. 2.1e-136;
Matches 336; Conservative 121; Mismatches 234; Indels 132; Gaps 18;
QY 1 MSNPGRRNGPVKRLVTLCAKMLVKQFFRLPDPFAKVVVDGSGQCHSTDVVKTLDPK 60
DB 1 MSNSAQSR---RIRVIVAADGLYKRDVFRFPDPAVLTVDGE-QTHTTTAKKTLPY 55
QY 61 WNOHYDLVIGKSDSVTISVNMHKKHKKQAGFLGCVRLLSNAINRLKDTGYORL--DLC 118
DB 56 WNETFVNVTDNSTIAIQVFDQKFP-KKKGQSGFLGVINLRVGDVLDLAIGGDEMLICDLK 114
QY 119 KLGNNDNDTVRGQIVVSLQ----- 137
DB 115 KL--NENTVVGKIIINLSTTAQLTLQVPSSAASGARTORTSITNDPOSSKSSVS RNA 172
QY 138 -----SRDRIGTGQVDCSRLPN-----DLPDGWEERTASGRIOYLNHITRTTQ 184
DB 173 SSRAGSTRONAPAASAPASEPTFSFEDQVGRLLPGWERTDNLGRYYVDHNRSTT 232
QY 185 WERTP-----RPASEYSPGRPLSCFVDENTPIGNTGATCGSSDPLAERRVRSORH 238
DB 233 WIRENLSSVAGAAAAAELHSSAS--SANVT-EGVQPSSSNAA-----RRTEASVLT--- 279
QY 239 RNYMSRTHLHTPPDLPEGVEQRTTQGGVFLHTQTGVSTWHDPRVPRDLSNIN----- 292
DB 280 ----SNATTAGSGELLFGWEQRYTPGGRFYFVDHNRITTTWVDRRQRIIRSYGGPNAT 335
QY 293 -----CBELGFLPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRQNLKDX 347
DB 336 IQQQPVSVQLGFLPSGWEMLTNTARVYFVDHNTKTTTWDPLPSSL-----DQ 384
QY 348 QQQQVSLCPDDETECLTVPRYKDLVQKILRQELSQ--QPOAGHCRIEVSREEIFEE 405
DB 385 -----NVQYKRDFFRKLIYF-----LSQPALHPLFPQCHIKVRRNHIFED 425
QY 406 SYROVMQWRPKDLKRLMIKFRGEGLDYGGVAREWLYLLEHMLNPPYGLFQYSRDDIY 465
DB 426 SYAEIMQSATDLKRLMIKFRGEGLDYGGVAREWLYLLEHMLNPPYGLFQYSRDDIY 485
QY 466 TLQINPDSAVNPEHLSYFHFVGRIMGMVAFHGHYIDGGFTLPPFYKQLLGKSIITLDMELV 525
DB 486 TLQINPHSGINPEHLSYFHFVGRIMGMVAFHGHYIDGGFTLPPFYKQLLGKSIITLDMELV 545
QY 526 DPDLHNSLWILENDITGVLDHTFCVEHNAYGEIIQHELKNGKSIPIVNEENKKEYVRLY 585
DB 546 DAERYSLWILDNDITGVLDHTFCVEHNAYGEIIQHELKNGKSIPIVNEENKKEYVRLY 605

586 VNWFLRGTEAFLALQGFNEVIOHLLKTFDEKELELLICGLKIDVNDKUNTELKH 645
 606 TWV-IQRIEOPNAPHEGSELIPQELINVFDEKELELLIGGSEIDMEDWKHKHYRS 664
 646 CTPDSNIVKFWKAVEFFEDEERRARLLQFVTGSSRPVQGFALQGAAGRLFTIHQIDA 705
 665 YSNDQIWKFWELMDWNSKSKLLQFTTGTGRIPVNGFKLQSGDGRKFTIEKAGE 724
 706 CTNNLPKATCNRRDIPYSEYKLYKLTALBETCGFAVE 748
 725 -PNKUPKATCNRLDLPYTSKOLDHKLSIAVEETIGFQOE 766

FLT 10

'3459

ABP73459 standard; Protein; 832 AA.

ABP73459;

30-JAN-2003 (first entry)

Candida albicans essential protein SEQ ID NO 7296.

Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 signal transduction; DNA replication; cell division; growth;
 proliferation; Candida albicans; fungicide; antifungal.

Candida albicans.

W0200253728-A2.

11-JUL-2002.

26-DEC-2001; 2001WO-US49486.

29-DEC-2000; 2000US-259128P.

20-FEB-2001; 2001US-0792024.

22-AUG-2001; 2001US-314050P.

(ELIT-) ELITRA PHARM INC.

Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

WPI; 2002-566694/60.

N-PSDB; ASZ32009.

Constructing strains for identifying gene products as effective targets
 for therapeutic intervention, by inactivating in the strain one allele
 of a gene and placing other allele of the gene under conditional
 expression -

Claim 44; SEQ ID NO 7296; 167pp + Sequence Listing; English.

The invention relates to constructing (M1) a strain of diploid fungal
 cells in which both alleles of a gene are modified, comprising modifying
 one allele by insertion or replacement by a cassette having an
 expressible selectable marker and modifying other allele by
 recombination, of a promoter replacement fragment with a heterologous
 promoter, so that expression of the second allele is regulated by the
 promoter. (M1) is useful for constructing a strain of diploid fungal
 cells in which both alleles of a gene are modified. The diploid fungal
 cells having both alleles modified are useful for identifying a gene that
 is essential to the survival or growth of a fungus, a gene that
 contributes to the virulence and/or pathogenicity of a fungus, a gene
 that contributes to the resistance of a diploid fungus to an antifungal
 agent, an antifungal agent that inhibits the growth of a diploid fungus
 and for identifying a therapeutic agent for treatment of a mammalian
 disease. (M1) is useful for identifying a compound which modulates the
 activity of a gene product, preferably enzymatic activity, carbon
 compound catabolism, biosynthetic, transporter, transcriptional,
 translational, signal transduction, DNA replication and cell division
 activity. The method is useful for identifying a compound having the
 ability to inhibit growth or proliferation of C. albicans cells and for

CC treating infection by C. albicans. The present sequence is that of an
 CC essential Candida albicans protein used in the method of the invention.
 CC Note: the sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.

XX

SQ Sequence 832 AA;

Query Match 37.9%; Score 1530.5; DB 23; Length 832;
 Best Local Similarity 38.4%; Pred. No. 5.7e-135;
 Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

6 RRRNGPVKRLTVLCAKLVKXDFRLLPDPFAKVVVDGSGQCHSTDTVKNLDPKKNQHY 65

9 RSNNTTINVKVAESLYKRDVFRQDPDFAVLTVDGS-QTKTTITAKKTLNPFYWNFTF 67

66 DLYIGKSDSV-TISVWNHKKLHKQAGFLGCVR-----LLSNAINRLKDTGYORLDLCK 119

68 N-FOAKEDSILVIQVFDQKP-KKQOQGLGVINVRIGDIVDLSNLSSEETITR--DLKK 123

120 LQPNNDTVRGQIVVSLQSRDRICTGGQVVD----- 150

124 --SNENLAVSCKIIIVI-SHNNSNGGGVTTATTGTGASSSNNIATITSGVNNLAIGSA 180

151 -----CSLFDN--- 157

181 TTANSTAQASDATVANGSGPTSSLPPIGQGHPEATAATPGGAAGAASQYSSFD 240

158 ---DLPGWBERRTASGRIQYLNIHTTTQWERPTRPASE----- 194

241 QYGRLLPPGWERTDNFGRTYVDNSRTTTWQPALHQSETERGQQQSQSEAEERQHRG 300

195 YSSPGR---PLSCFVDENTPISGTNGATCGQSS-----DPLAERRVRSQHRN 240

301 RPLPGEGSVSL-----PTSGNSITSGNVTNAGANTFVNPAAYVMAASGATTS 352

241 YMSRTHLTPDPEGYEORTTQOQVYFLHTQTGVSTWHDPRVPRDLNSIN----- 292

353 GLG-----ELPSGWEQRTTEGRPYFVDHNTRTTTWDPDPRQVIRTFGNNTIQQQ 404

293 -CEELGPLPPGWEIRNTATGRVYFVDNNRTTQTDPRLSANLHLVNROKLDKQOQQQ 351

405 PVSQGLPLPSGWEMRLNTARVYFVDHNTKTTTWDPRLPSSL-----DQ--- 449

352 VVSLCPDDTECLTVPRYKRDVOKILKRLQELSQOQPQ---AGHCRIEVSREIPEESY 407

450 -----NVQYARDPRKVIYFR-----SQPALRILPQOCHIKVRRDHIFEDSY 492

408 RQVMKMRPKDLMKRLMIKFERGEGLDYGVAREWLYLLSHMLNPPYGLFOYSRDDIYTL 467

493 QEIMRQTPEDLKKRLMIKFERGEGLDYGVAREWLYLLSHMLNPPYGLFOYSRDDIYTL 552

468 QINPDSAVNPEHLSYFHFVGRIMGMVPHGVHIDGGFTLPPYKOLLGKSLTDDMELVDP 527

553 QINPDSAVNPEHLSYFHFVGRIMGMVPHGVHIDGGFTLPPYKOLLGKSLTDDMELVDP 612

528 DLHNSIVATLENDITGVLDHTFCVEHNAIGEIIQHELKPNKSIPIVNEENKKEYVRLVYN 587

613 EFVRSKLWLDNDITGILDLTFSAEESFGEIVVDLKPGRDIEVTEENKHEYVELITE 672

588 WFLRGIEAQFALQKQNEVUPHLLKTFDEKELELLICGLKIDVNDKUNTELKHCT 647

673 WRISKVEVEQFKAFTDGNELIPQELVNVFDERLELLIGGLAIEDCEDWKHKHYDYGYQ 732

648 PSNIVKFWKAVEFFEDEERRARLLQFVTGSSRPVQGFALQGAAGRLFTIHQIDA 707

733 ENDQVQLQWFKWCKINWSDSEQKALLQFTTGTSRIPVNGFKDQSGDGRPRFTIEKAGE-A 791

708 NNLPKAHTCFNRIDIPYSEYKLYKLTALBETCGFAVE 748

792 NQLPKSHTCFNRVDLPYTDYESLAKQLTLAVEETVGFGQE 832

ILT 11	QY	11	PVKLR-L-TVLCAKNLVKDPRF-----LDPFAKVVVDGSGOCH	48
6104	DB	15	PASSRLCRVIAADGLYXRDRFRKSYILAILLVRLITSGAFAGFPDFPFAVATVGBE-QTH	73
ABU26104;	QY	49	STDVYQNTLDPKMNQHYDLYIGKSDSVTISWNHKKIHKQKAGFLGCVRL--	100
16-APR-2003 (first entry)	DB	74	TTSVIKKTLPNPNWEMFDRWNEDSILAIQIFDQKXF-KKQDQFLGVINVRIGDVIDLQ	132
Aspergillus fumigatus essential gene protein #762.	QY	101	-----SNAINRLK-----DTGYRDLDLKLPNDNDTVRGQIVSL-----	136
Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection; cancer; contamination; biofilm; antibody; immune response.	DB	133	MGDGESLPINHSQDSRLTFRLLLETEMLTDLKK--SNDNLVVHGLIILNLTSLST	190
Aspergillus fumigatus.	QY	137	-----QSRDRIGTGQVVDCS-----	152
WO200286090-A2.	DB	191	PNTQANGLHRSHVQSSTSSGLVFQVAPSSSHPAASGTAPVDPASNPSPINQPVSTTR	250
31-OCT-2002.	QY	153	-----RUFND-----LDPWEERTASGRIOYLNHITRTTQ	184
23-APR-2002; 2002WO-US13142.	DB	251	PSTAAPASAAAGAAVNSHGSRTNLSFSDSQRLPAGWERREDNLGRTYVDHNTRTTT	310
23-APR-2001; 2001US-285697P.	QY	185	WERPTRPASEYSSPGRPLSCFVDENTPISGTNGATCGQSDPDLAERRVRSORHRYM--	242
27-APR-2001; 2001US-287066P.	DB	311	W---TRPSSNYNEHAQ-----RSQREANMQLERRAHOQSRMLP	344
05-JUN-2001; 2001US-295890P.	QY	243	-----SRTHLHTPP-----DLPEGYEQRTTQOQVY	268
09-JUL-2001; 2001US-303899P.	DB	345	EDRTGANSPLPSSQAHTPPAGGGSANAVSMMATGATTAGTGELPFGWEQRTTPEGREY	404
31-AUG-2001; 2001US-316362P.	QY	269	FLHTQTGVSTWHDPRVPRDL-----SNIN-----CEELGPLPGWEIRNTATGRVY	314
(BLIT-) ELITRA PHARM INC.	DB	405	FVDENRTTTTWDPRROQYRMYGONANGNTTIOQOPVSQLGFLPSGWEMLTNTARY	464
Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;	QY	315	FVDHNNRTTQFTDPRLSANLHLVLRNQNLKQDQOQVWSLCPDDTECLTVPRYKDLVQ	374
WPI; 2003-093124/08.	DB	465	FVDHNTKTTTWDPRLPSSL-----DQ-----GVPOYKDFRR	497
New purified or isolated nucleic acids of essential genes of	QY	375	KLKILRQELSQQP-----QAGHCEIEVSREEIESYRVQVMKRPKDLMKELMKERGE	430
Aspergillus fumigatus, useful for treating or preventing infections by	DB	498	KLIVFR-----SQPALRMGQCHVKVRNNIFEDSYAEMROSADLKKRLMKFDGSD	552
A. fumigatus, or for treating a non-infectious disease in a subject	QY	431	GLDYGVAREWLYLLSHEMLNPPYVGLFOYGRDDIYTLQINPDVANPPEHLSYFHFVGRIM	490
e.g. cancer	DB	553	GLDYGLSGEFFFLLSHEMFNPFVCLFEYSADHNYTLQINPHSGVNPHEHLYFKFGRVV	612
Disclosure; Page -; 175pp; English.	QY	491	GNAPVGHYIDGGTLPFPYKQLGKSLTLDMLVDPDLHNSLVWLENDITGVLDHTFC	550
The invention relates to novel purified or isolated nucleic acids of	DB	613	GLAIFHRFLDSPFIFAGFYKMLRKKVSLQDMGVEDLHRLNTWTWMDNIEGVLELTF	672
essential genes of Aspergillus fumigatus. The isolated nucleic acids of	QY	551	VEHNAYGRIIQHELKPNKGKIPVNEENKKEYVLYVNRFLRGIEAQFLALQKGFNEVIP	610
the invention are used to treat or prevent infections by a pathogenic	DB	673	VDEKFGERRIDILKPGGRDIPVTNENKAEYVLTVMKIVKEVEQFNAMFGFNELIP	732
organism such as A. fumigatus, to treat a non-infectious disease in a	QY	611	QHLKTPDEKELELIICGLKIDVNDKVNTRLKHCTPDPSNIVKFWKAVEFPDEBRRAR	670
subject (e.g. cancer), to prevent or contain contamination of an object	DB	733	ADLVNFDERELELLIGIADIDVDDWKHKTGYRGYQESDEVQNFVKIVRSWDAEQKR	792
by A. fumigatus, or to prevent or inhibit formation on a surface of a	QY	671	LLQFTGSSRVPLOGKALOGAAGPRFLTTHQIDACTNMLPKAHTCFNRIDIPPYESYEK	730
biofilm comprising A. fumigatus. The polynucleotides are useful for	DB	793	LLQFTGTSRIPVNGFKLOSGDGPFRFTIEK--SGDPAALPKSHTGFNRLDLPYKSJET	851
expressing recombinant protein for characterisation, screening or	QY	731	LYEKLLTAIBETCGFAVE 748	
therapeutic use, as markers for host tissues in which the pathogenic	DB	852	LEHKMSIAVEETLFGQGE 869	
organisms invade or reside, for comparing with the DNA sequence of A.				
fumigatus to identify duplicated genes or paralogues having the same or				
similar biochemical activity and/or function, for comparing with DNA				
sequences of other related or distant pathogenic organisms to identify				
potential orthologous essential or virulence genes, for selecting and				
making oligomers for attachment to a nucleic acid array for examination				
of expression patterns, for raising anti-protein antibodies, as an				
antigen to raise anti-DNA antibodies or to elicit another immune				
response, and for identifying polynucleotides encoding the other protein				
with which binding occurs or to identify inhibitors of the binding				
interaction. The polypeptides may be used to raise antibodies or to				
elicit immune response, as a reagent in assays designed to quantitatively				
determine levels of the protein in biological fluids, as a marker for				
host tissues in which pathogenic organism invade or reside, and to				
isolate correlative receptors or ligands in the case or virulence				
factors. This sequence represents a protein of one of the essential genes				
of Aspergillus fumigatus of the invention.				
Sequence 869 AA;				
ery Match				
st Local Similarity 36.2%; Pred. No. 1.2e-127;				
atches 332; Conservative 114; Mismatches 225; Indels 243; Gaps 21;				

RESULT 12
AAW93167
ID AAW93167 standard; Protein; 975 AA.
XX
AC AAW93167;
XX
DT 24-MAY-1999 (first entry)
XX

269 FRSRRISEDLPEPSEGDPWET-ISEEVNIAGDSLGVLPFPPASPGSRTSPQEL 327

```
PT      domain  
CIT     /note= "claimed (claim 90) HECT domain"
```

WO9737223-A1.
09-OCT-1997.
03-APR-1997; 97WO-US05547.
03-APR-1996; 96US-0630916.
(CYTO-) CYTOGEN CORP.
(UNIC-) UNIV NORTH CAROLINA.
Fowlkes DM, Kay BK, Pirozzi G;
WPI; 1997-503234/46.
N-PSDB; AAT95700.
Identifying cell signalling and growth regulatory polypeptides by
reaction with multivalent recognition complex - polypeptides are
useful in targetted drug selection
Claim 48; Fig 23; 220pp; English.
The present sequence represents a novel protein WWP4. The WWP4 gene was
identified and isolated from a cDNA expression library generated from
LNCap prostate cancer cell line, using peptides AAW38063-64. These
peptide recognition units are based on the sequences of WW domain binding
domains of the alpha and gamma subunits of epithelial sodium channel
protein. The WW domain is a small functional domain found in a large
number of proteins from a variety of species including humans, nematodes
and yeast. Its name is derived from the observation that two tryptophan
residues, one in the amino terminal portion of the WW domain and one in
the carboxyl terminal portion, are conserved. Most proteins containing
WW domains have a function involving cell signalling and growth
regulation or the organisation of the cytoskeleton. Polypeptides
containing a WW domain are identified by treating a multivalent
recognition unit complex that has selective binding affinity for a WW
domain, with many polypeptides and identifying those with selective
affinity for the complex. Proteins containing WW domains are used for
targeted drug screening, i.e. to identify potential modulators of
specific WW domain interactions. The valency of the recognition unit is
important in determining specificity of interaction with WW domains. In
multivalent form specificity is relaxed, but not lost, so proteins
containing WW domains similar, but not identical, to the sequence of the
peptides' target WW can be detected, including new polypeptides.
Sequence 724 AA;
Very Match 33.6%; Score 1355.5; DB 18; Length 724;
1st Local Similarity 43.7%; Pred. No. 1.6e-118;
Itches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;
137 QSDRIGTGQGVVDCGRLEPDN--DLPGWEERTASGRIQVLAHITRTTOWERPTPAS 193
111 RARSSTVTGGEPTPSVAVVHTTGLPSGWEERKDAKGRYYVNNHNRITTTWTRIMQLA 170
194 EYSSPG-----RPLSC---FYDENTPISTNGATCGSSDRLRAERVR 235
171 EDGASSATNSNNHLEIPRIAPRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPSPQ 230
236 QRHRNYSRTHLTPPDLPPEGVEQRTQGGVYFLHTQTGVSTWHDPRV-----RDLN 290
231 PPSYNSPKQHKVTKQFLPPGWEIRIAPNGRPFIDHNTKTITWEDPLKPPVHRSKTS 290
291 INCEELGPPPGWEIRIATGAVYFVDHNRRTQTDPRLSANLHLVLRNQNLKDDQQQ 350
291 LNPDNLGPPPGWEIRIHLDGRTFYDHNKTIQWEDPRL-----QN----- 332
351 QVYSLCPDDTECLTVPR--YKRDVLQKLKILKQLSQQQPGACHRIEVSREEIFEESYR 408
333 -----PAITGPAVYSREFKQVDYPRKLLKPADIPNRFEMKLNHNIFEESYR 382
409 QVMKRPKDLWK-RLMKRFGEGLDYGGVAREWLYLLSHENLAPFYGLFQYSRDDIYTL 467

Db 383 RINSVKPDPVLKARLWIEFESEKGLDYGVAWEWFLLSKEMFNPYGLFEYSATNTYL 442
Qy 468 QINPDSAV-NPEHLSYFHFVGRIMGMVAFHGHVIDGGFTLPYKQLLGKSIITLDDMELVD 526
Db 443 QINPNSGLCNEDHLSYFTFIRKGVAGLAVFHGKLLDGFIRPPYKMWMLGKQITLNDMESVD 502
Qy 527 PDLNSLVWILENDITGVLDHTFCVEHNAYGEITIQHELKPNKSGSIIPVNEENKKEYVRLVX 586
Db 503 SEYNSLKWILENDPT-ELDMFCIDEENFGQTYQVDLAPNGSEINVTNENKREYIDLVI 561
Qy 587 NWRFLRGIEBAQFLALQKGFNEVIPQHLKTFDFKELELIICGLGKIDVNDKWNTRLK-- 644
Db 562 QMRFVNRVQKMAFLGEGFTLELLPIDLIKIFDENELELLMCGLDVNDVNDWRQHSYKNG 621
Qy 645 HCTPDSNIVKWFKAWEFFDEERRARLLQPTVGSSVPLQGFKALQGAAGPRLFTTHQID 704
Db 622 YC-FNHVPIQWFAVLLMDAKRIRLLQVTGTSVPMNGFAELYGSGNGPQLFTIEQWG 680
Qy 705 ACTNNLPKATCFNRIIDIPPYSEYKLYEKLTAIBETCGF 745
Db 681 S-PEKLPKATCFNRLDLPYETFEDLQEKLLMAVENAQGF 720
RESULT 14
AA25170
ID AAY25170 standard; Protein; 927 AA.
XX
AC AAY25170;
XX
DT 09-SEP-1999 (first entry)
XX
DE Human KIAAN ligase protein fragment.
XX
KW Inhibitor; ubiquitin-mediated proteolysis; IkappaB polypeptide; HECT;
KW ubiquitin-conjugating system; Homologous to E6-Ap carboxyl terminus;
KW ligase; ubiquitin; ubiquitination; RSC ligase; KIAAN ligase; UBC4;
KW MAD; modulator; treatment; proliferative disorder; apoptosis; sepsis;
KW differentiative disorder; viral infection; tissue wasting disorder;
KW cachexia; malignancy; inflammatory disease; parasitic disease;
KW tuberculosis; IL-2 therapy; rheumatoid arthritis; rheumatoid disease;
KW osteoarthritis; gouty arthritis; respiratory distress syndrome;
KW cerebral malaria; chronic pulmonary inflammatory disease; silicosis;
KW pulmonary sarcoidosis; bone resorption disease; reperfusion injury;
KW Graft versus host reaction; allograft rejection; Crohn's disease;
KW ulcerative colitis; pyresis; multiple sclerosis; autoimmune diabetes;
KW systemic lupus erythematosus; leprosy; AIDS.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..927 /note= "partial sequence, no start codon given"
XX
PN WO9904033-A1.
XX
PD 28-JAN-1999.
XX
PF 16-JUL-1998; 98WO-US14638.
XX
PR 16-JUL-1997; 97US-0895601.
XX
PA (MITO-) MITOTIX INC.
XX
PI Beer-Romero P, Glass SJ, Rolfe M, Strack PR;
XX
DR WPI; 1999-132274/11.
DR N-PSDB; AAX78494.
XX
PT Identifying modulators of I-kappa-B proteolysis - used to develop
PT products for treating e.g. proliferative and/or differentiative
PT disorders, infections, tissue wasting, cachexia or AIDS
XX

Disclosure; Page 68-71; 79pp; English.

This invention describes a novel assay identifying an inhibitor of ubiquitin-mediated proteolysis of a IkappaB polypeptide. The method comprises: (a) providing a ubiquitin-conjugating system including the IkappaB polypeptide and a HECT (Homologous to E6-Ap Carboxyl Terminus) ligase and ubiquitin, to promote ubiquitination of the IkappaB polypeptide by the HECT ligase; (b) contacting the ubiquitin-conjugating system with a candidate agent; (c) measuring a level of ubiquitination of the polypeptide in the presence of the candidate agent and (d) comparing the measured level of ubiquitination in the presence of the candidate agent with ubiquitination of the IkappaB polypeptide in the absence of the candidate agent; where a statistically significant decrease in ubiquitination of the IkappaB polypeptide in the presence of the candidate agent is indicative of an inhibitor of ubiquitination of the IkappaB polypeptide. The assays can be used to identify compounds which modulate binding and/or ubiquitination of an IkappaB (or other cellular or viral substrate) by a HECT ligase, such as RSC or KIAAN. Such modulators can be used e.g. in the treatment of proliferative and/or differentiative disorders, to modulate apoptosis, in the treatment of viral infections, and in the treatment of tissue wasting disorders e.g. cachexia secondary to infection or malignancy, cachexia secondary to human AIDS, inflammatory diseases, parasitic diseases, tuberculosis and high dose IL-2 therapy; rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis and other arthritic conditions, sepsis, respiratory distress syndrome, cerebral malaria, chronic pulmonary inflammatory disease, silicosis, pulmonary sarcoidosis, bone resorption diseases, reperfusion injury, graft versus host reaction, allograft rejections, Crohn's disease, ulcerative colitis, or pyresis, in addition to a number of autoimmune diseases such as multiple sclerosis, autoimmune diabetes, systemic lupus erythematosus, and ENL in leprosy, HIV, and AIDS. This sequence represents a human KIAAN ligase which is used in the method of the invention.

Sequence 927 AA;

ary Match 33.5%; Score 1354.5; DB 20; Length 927;
 at Local Similarity 34.0%; Pred. No. 3e-118;
 tches 329; Conservative 123; Mismatches 232; Indels 285; Gaps 28;
 4 PGRNRGPNV-----KRLTLVLCARNLVKKDFRLLPDPFAK 38
 13 PGRGRAPVCKRSEMATCAVEVFGLLEDENSRIVRVRIAGIGLAKKIDLCASDPYVR 72
 39 VV-----VDSGGOCHSTDTVKTLDPKN-----OH-----YDLVIGK 71
 73 VTLDPMMGVLTSTVQTKTKSLNPKMNEELFRVHPQOHRLLFEVDENRLRDPDLQ 132
 72 SD--SVTISVMN-----HKIKHKQAGFLGCVRLLSNAINRLKDTGYQRL 115
 133 VDVPVLYPLPTENPLRERYTKDFVLPKSHKSVKGYL-----RLKWT----- 176
 116 DLCKLGNPDNTRVGVQVVSQSRDRITGSGQVVD-----C---SRLFNDLDPGHEER 167
 177 YLPKTSGSDD-----NAEQAELEFGWVLDQPDAAACHLQQQOQSPPLPGWEERQ 228
 168 TASGRIOYLNHLTRITQWERPT-----RPASE----- 194
 229 DILGRYYVNHESRTQWKRPQDNLDAENGNILOAQPAFTTRQISEETESVDNQE 288
 195 -----YSS---PGRFLSCFVDNTPISGT-----NGATCGQSSDPR 227
 289 SSENWEIIEDEATYSSQAFPPSPSSNLDVPHLAEELNARLTIFGNSAVSPASSN 348
 228 LAERR-----VRSQRHNYM-----SRTHLHTPPD- 252
 349 HSRRRGSLQAYTFESQPTLPLVLLTSSGLPPGWEEKQDERGRSYVDHNRITTTWKPTV 408
 253 -----LPEGYEQRTQQQGVYFLH 271
 409 QATVETSLTSSQSSAGPQSQASTSDSGQVQTPQSEIEQGLPKGWVHRAPNGRPFID 468
 272 TQGVSTWHDPR--VPR-----DLNINCEELGPLPGWEIRNTATGRVYFVDHNR 321

Db 469 HNTKTTWEDFLKTPAHLRGKTSLDTSN-----DLGPLPGWEERTHTDGRIFVINENIK 524
 Qy 322 TQQTDPRLSANHLVLRQNLKDDQQOQVSLCPDDTECLTVPR--YKRDVLQVKLIL 379
 Db 525 RTQWEDPRL-----ENVAITGPAVPYGRDYKRYEYFF 556
 Qy 380 ROELSQQQPAQCHRIEVSREIPEESYRQVMKRPKDLWK--RLMIKFRGEBGLDYGVA 438
 Db 557 RRLKKQNDIPNKFEMKLRATVLEDSYRIMGVKRADPLKARLWIEFDGKGLDYGVA 616
 Qy 439 REMWYLLSHEMLNPYGLFQVSRDDIYTLQNPDSAV--NPEHLSYFHFVGRIMGVAFHG 497
 Db 617 REWFFLISKEMFNYYGLFEYSATDNYTLQINPNSGLCNEHLSYFHFVGRIMGVAFHG 676
 Qy 498 HYIDGGFTLPFFYKOLLGKSITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYG 557
 Db 677 KLIDGFFIFFYMMHLKRPITLHDSVDSVYNSLRWILENDPT-ELDLRFIIDEELFG 735
 Qy 558 EITQHELKNGKSIPIVNEENKKEYVRLVYNNRFLRGIEAQFLAQKNGNEVIPHLLKTF 617
 Db 736 QTHQHELKNGSGSEIVVTKNKKKEYILVQWRFNRIQKMAAFKEGFFELIPQDLIKIF 795
 Qy 618 DEKELELITCGLGKIDVNDKYNTRIKH--CTPDSNIVKFWKAVEFFDEERRARLLQFVT 676
 Db 796 DENELELLMCGLGVDVNDWRHRTKFKNGYSANHQVIOFWKAVLWMDSEKRIALLQFVT 855
 Qy 677 GSSRVLPQGFALQGAAGPRLFTIHQIDACTNNLPKANTCNRIDIPYSEYKLYEKL 736
 Db 856 GTSRVPWNGFAELYGSGNGPSQFTVEQ--GTPEKLPRAHTCFNRLDLPYSEFEELMDKLQ 914
 Qy 737 TAIEETCGF 745
 Db 915 MAIENTQGF 923
 RESULT 15
 AAEE32725
 ID AAEE32725 standard; Protein; 927 AA.
 XX
 AC AAEE32725;
 XX
 DT 24-MAR-2003 (first entry)
 XX
 DE Human NEED4 long form protein.
 XX
 KW Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis; poliomyelitis; HIV; measles; protein therapy; human; NEED4.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 381..406
 FT notes="WW domain"
 FT Domain 821..923
 FT notes="HECT domain"
 XX
 PN WO200290549-A2.
 XX
 PD 14-NOV-2002.
 XX
 PF 12-MAR-2002; 2002WO-IB02106.
 XX
 PR 12-MAR-2001; 2001US-275224P.
 PR 31-JUL-2001; 2001US-308958P.
 PR 07-DEC-2001; 2001US-340170P.
 XX
 PA (PROT-) PROTOLOGICS LTD.
 XX
 PI Greener T, Moskowitz H, Reiss Y, Alroy I;
 DR WPI; 2003-111976/10.
 DR N-PSDB; AAD50459.

New protein complex comprising HECT-RCCL1, viral maturation scaffolding protein (VNSP), and/or HIV gag protein, useful for treating viral infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles, or Ebola -

Disclosure; Fig 15; 150pp; English.

The invention relates to a method for modulation of viral maturation. The invention also provides an isolated protein complex comprising a HECT-RCCL1 selected from HECT-NW, HECT-RCCL1, Gag protein, Gag late domain, P13, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, GPase, E2 enzyme, teg101, cullin, HRC1, HRC2, HRC3, Nedc4 -like protein or clathrin. The complexes, proteins, antibodies and methods are useful for treating viral infections, such as lymphosarcoma, human immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola and for inhibiting budding in a subject. They are also useful in diagnostic assays for determining whether a cell is infected with a virus and for characterising the nature, progression and/or infectivity of the infection. The invention is also useful in protein therapy. The present sequence is human NEDB4 protein used to illustrate the method of the invention.

Sequence 927 AA;

ery Match 33.5%; Score 1354.5; DB 24; Length 927;

st Local Similarity 34.0%; Pred. NO. 3e-118;

tches 329; Conservative 123; Mismatches 232; Indels 285; Gaps 28;

4 PGRRRNGPV-----KRLTLVCAKNLVKDFRLLPDPFAK 38

13 PGRGRARPVCKRSEMATCAVEVGLLEDENSIRVRVIAGIGLAKKILGASDPYVR 72

39 VV-----VDGGGCHSTDTKNTLPKMN-----QH-----YDLYIGK 71

73 VLYDPMNGVLTSTVQTKIKKSLNPKWNEELLFRVHPQORLLFEVPDENRLTRDDFLGQ 132

72 SD--SVTISVKN-----HKIHKQAGAGFLGCVRLLSNAINELKDTGYQRL 115

133 VDVPVLYPTENPLRERYTFKDFVLHPRSHKSVKGYL-----RLKWT----- 176

116 DLCKLGPNDNDTVRGQIVVSLQSRDRIQTGGQVVD-----C-----SRLFNDNDLPDGWEERR 167

177 YLPKTSGEDD-----NAEQAEELPFGVVDQPDAAACHLQQQQQEPSPPLPGWEERQ 228

168 TASGRIOVLNHTTRTQWERT-----RPAE----- 194

229 DILGRYYYNHESRETOKRPTQDNLTDAENGNIQLAQRAFTTRQISETESVDNQE 288

195 -----YSS-----PGRPLSCFVDENTPISGT-----NGATCGQSSDPR 227

289 SSENWEIIRDEATWYSSQAFSPSPSSNLDVPHLABELNARLTIFGNSAVSQPASSN 348

228 LAER-----VRSQRHNYM-----SRTHLHTFPD- 252

349 HSSRRGSLQAYTFEEQPTLPVLLPTSSGLPPGWEKQDERGRSYVDHNSRTTWTKPTV 408

253 -----LPEGVEORTTQGGQVVELH 271

409 QAVTSTQTSQSAGQSASTSDSCQQTQSEIEQGFPLKGVNVRHAPNGRPFPPID 468

272 TGTGVSTWHDPR--VPR-----DLNINCEBELGPLPGWEIRNTATGRVYFVDHNNR 321

469 HNTKTTTWTEDPLKIPAHLRGKTSLDTSN-----DLGPLPGWEERTHTDGRIFVYNHNK 524

322 TTCTPTDPLSANHLVLRNQKLDQQQQQVSLCPDDTECLTVPR--YKEDLVQKLIKIL 379

525 RTQWEDPRL-----ENVAITGPAVYGRDYKRYEYFF 556

380 ROELSQOQFQAGCHRIEVSREIIEESYRQVMKRPKDLWK-RLMIKFRGEEGLDYGGVA 438

557 RRLKQKQNDIPNKFEMKLRRATVLEDSYRRIMGVGRADFLKARLWIEFDGEGKLDYGGVA 616

QY 439 REMLYLLSHMLNPYYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFYGRIMGMAVEFHG 497

Db 617 REWFFLLSKEMFPYYGLFEYSATDNYTLQINPNSGLCNEDHLSYFKFGRVAGVAVYHG 676

QY 498 HYIDGGTLPFYKQLIGKSTLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNNAVY 557

Db 677 KLDDGFFIRPFPYKMLHKKPITLHDMESVDSEYNSLRWILENDPT-ELDLRFIDDEELFG 735

QY 558 BIIQHEIKPKNGKSPVNEENKKEYVRLVYNWRFLRGIEAQFLALOKGNEVIPAQLLKTFF 617

Db 736 QTHQHELVNGSGSELVTVNKNKKEYIVLVQWRFVNRIOQMAAFKEGFFELIPQDLIKIF 795

QY 618 DEKELELIICGLGKIDVNDKVNTRLKH-CTPDSNIVKFWKAVFEFFDEERARLLQFVT 676

Db 796 DENELELLMCGLDGVDVNDWREHTKYKNGYSANHQVQWFKAVLMDSEKIRLLIQFVT 855

QY 677 GSSRVPLQGFKALQGAAGPELFTIHQIDACTNNLPAKHTCENRIDIPPVYESYEKLYEKL 736

Db 856 GTSRVPNGFAELYGSGNGSQSFVTEQW-GTPEKLPRAHTCFNRLLDLPPIYESFEELWDLQ 914

QY 737 TAIBETCGF 745

Db 915 MAIENTQGF 923

Search completed: February 20, 2004, 15:28:33

Job time : 45.6968 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
protein - protein search, using sw model
n on: February 20, 2004, 15:26:11 ; Search time 15.7634 Seconds
(without alignments)
2007.718 Million cell updates/sec
tles: US-10-009-945-4
xfect score: 4038
quence: 1 MSNPGRRNGPVKRLTLVLC.....EKLYKLTAIEETCGFAVE 748
oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
arched: 328717 seqs, 42310858 residues

tal number of hits satisfying chosen parameters: 328717

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

tabase : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pdp:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pdp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match %	Length DB	ID	Description
1	3939.5	97.6	735	3	US-08-539-205A-2
2	3939.5	97.6	735	4	US-09-392-163A-2
3	1564	38.7	766	4	US-08-539-205A-4
4	1564	38.7	766	4	US-09-392-163A-4
5	1354.5	33.5	927	3	US-08-895-601-6
6	1353	33.5	834	3	US-08-539-205A-6
7	1353	33.5	834	4	US-09-392-163A-6
8	1336	33.1	854	2	US-09-070-060-4
9	1336	33.1	854	3	US-09-357-746-4
10	1329	32.9	852	2	US-09-070-060-3
11	1329	32.9	852	3	US-09-357-746-3
12	1298.5	32.2	906	3	US-08-630-916A-48
13	1134.5	28.1	683	3	US-08-630-916A-46
14	534.5	13.2	874	2	US-08-247-904B-8
15	534.5	13.2	874	3	US-08-767-942A-21
16	513	12.7	866	1	US-08-100-692-1
17	513	12.7	866	2	US-08-674-030-1
18	450.5	11.2	1083	3	US-08-895-601-5
19	219	5.4	486	3	US-08-348-518C-2
20	207	5.1	472	3	US-08-348-518C-5
21	207	5.1	472	3	US-08-476-509B-5
22	162.5	4.0	54	3	US-08-630-916A-124
23	158	3.9	448	3	US-08-476-509B-2
24	156.5	3.9	55	3	US-08-630-916A-75
25	154.5	3.8	454	3	US-08-348-518C-4
26	154.5	3.8	454	3	US-08-476-509B-4
27	153	3.8	224	3	US-08-630-916A-50

28	150	3.7	51	3	US-08-630-916A-117	Sequence 117, App
29	147.5	3.7	54	3	US-08-630-916A-74	Sequence 74, Appl
30	146	3.6	51	3	US-08-630-916A-73	Sequence 73, Appl
31	145.5	3.6	58	3	US-08-630-916A-84	Sequence 84, Appl
32	141	3.5	54	3	US-08-630-916A-118	Sequence 118, App
33	140	3.5	38	3	US-08-630-916A-26	Sequence 26, Appl
34	140	3.5	38	3	US-08-348-518C-18	Sequence 18, Appl
35	140	3.5	38	3	US-08-476-509B-18	Sequence 18, Appl
36	136	3.4	38	3	US-08-630-916A-36	Sequence 36, Appl
37	135	3.3	51	3	US-08-630-916A-116	Sequence 116, App
38	132	3.3	3	3	US-08-630-916A-32	Sequence 32, Appl
39	129.5	3.2	1105	2	US-09-220-157A-2	Sequence 2, Appli
40	129.5	3.2	1105	4	US-08-710-249-2	Sequence 2, Appli
41	127.5	3.2	335	2	US-08-844-312-2	Sequence 2, Appli
42	127	3.1	51	3	US-08-630-916A-115	Sequence 115, App
43	124	3.1	38	3	US-08-630-916A-24	Sequence 24, Appl
44	124	3.1	38	3	US-08-348-518C-15	Sequence 15, Appl
45	124	3.1	38	3	US-08-476-509B-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-539-205A-2
; Sequence 2, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Neifsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-2

Query Match	97.6%	Score	3939.5	DB	3	Length	735
Best Local Similarity	99.0%	Pred. No. 0					
Matches	733	Conservative	0	Mismatches	2	Indels	13
Gaps	1						
QY	1	MSNPGRRNGPVKRLTLVLC	AKNLRVKKDFRLLPDPFAKVVVDGSGGCHSTDIVKNTLDPK	60			
Db	1	MSNPGRRNGPVKRLTLT	-----GLDPFAKVVVDGSGGCHSTDIVKNTLDPK	47			
QY	61	WNQHYDIYIGKSDSVIISVWVNHKKIKKQAGPLGCVRLLSNAINRLKDTGYORLDLCKL	120				

48 WNOHYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 107
121 GNDNDYVRGQIWSLQSRDRIGTGGVVDCSRFPDNDLPGWEERTASGRIOYLNHIT 180
108 GNDNDYVRGQIWSLQSRDRIGTGGVVDCSRFPDNDLPGWEERTASGRIOYLNHIT 167
181 RTQWERPTRPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 240
168 RTQWERPTRPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 227
241 YMSRTHLTPDLPPEGYEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCEELGPLP 300
228 YMSRTHLTPDLPPEGYEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCEELGPLP 287
301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDQGGVYSLCPDDT 360
288 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDQGGVYSLCPDDT 347
361 ECLTVPRYKRDVLVQKILRQELSQQPGAGHCRIEVSREEIFEEYSYQVMKRPKDLWK 420
348 ECLTVPRYKRDVLVQKILRQELSQQPGAGHCRIEVSREEIFEEYSYQVMKRPKDLWK 407
421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
408 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 467
481 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 540
468 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 527
541 ITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLVYNNWRLRGIEAQFLA 600
528 ITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLVYNNWRLRGIEAQFLA 587
601 LQKGFNEVIFQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
588 LQKGFNEVIFQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 647
661 EFPDEERRARLLOFTVGTSSRVPLOQFKALQGAAPRLFTTHQIDACTNNLPKATCFNRI 720
648 EFPDEERRARLLOFTVGTSSRVPLOQFKALQGAAPRLFTTHQIDACTNNLPKATCFNRI 707
721 DIPPEYSYKLYEKLTAETTCGFAVE 748
708 DIPPEYSYKLYEKLTAETTCGFAVE 735

MULT 2

09-392-163A-2
Sequence 2, Application US/09392163A
Patent No. 6503742
GENERAL INFORMATION:

APPLICANT: Beach, David H.

APPLICANT: Caligiuri, Maureen

APPLICANT: Neisky, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELLIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 735 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-392-163A-2

Query Match 97.6%; Score 3939.5; DB 4; Length 735;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;

QY 1 MSNFGRRNGPVKRLTLVCAKLVKKDFRLLPDPFAKVVVDGSGGCHSTDTVKNTLDPK 60
Db 1 MSNFGRRNGPVKRLTLVCAKLVKKDFRLLPDPFAKVVVDGSGGCHSTDTVKNTLDPK 47
QY 61 WNOHYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 120
Db WNOHYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 107
QY 121 GPNDNDYVRGQIWSLQSRDRIGTGGVVDCSRFPDNDLPGWEERTASGRIOYLNHIT 180
Db GPNDNDYVRGQIWSLQSRDRIGTGGVVDCSRFPDNDLPGWEERTASGRIOYLNHIT 167
QY 181 RTTOWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 240
Db RTTOWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 227
QY 241 YMSRTHLTPDLPPEGYEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCEELGPLP 300
Db YMSRTHLTPDLPPEGYEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCEELGPLP 287
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDQGGVYSLCPDDT 360
Db PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDQGGVYSLCPDDT 347
QY 361 ECLTVPRYKRDVLVQKILRQELSQQPGAGHCRIEVSREEIFEEYSYQVMKRPKDLWK 420
Db ECLTVPRYKRDVLVQKILRQELSQQPGAGHCRIEVSREEIFEEYSYQVMKRPKDLWK 407
QY 421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 467
QY 481 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 540
Db SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 527
QY 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLVYNNWRLRGIEAQFLA 600
Db ITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLVYNNWRLRGIEAQFLA 587
QY 601 LQKGFNEVIFQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Db LQKGFNEVIFQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 647
QY 661 EFPDEERRARLLOFTVGTSSRVPLOQFKALQGAAPRLFTTHQIDACTNNLPKATCFNRI 720
Db EFPDEERRARLLOFTVGTSSRVPLOQFKALQGAAPRLFTTHQIDACTNNLPKATCFNRI 707
QY 721 DIPPEYSYKLYEKLTAETTCGFAVE 748
Db DIPPEYSYKLYEKLTAETTCGFAVE 735

SU1T 3
 -08-539-205A-4
 Sequence 4, Application US/08539205A
 Patent No. 6001619
 GENERAL INFORMATION:
 APPLICANT: Beach, David H.
 APPLICANT: Caligiuri, Maureen
 APPLICANT: Nefsky, Bradley
 TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/539,205A
 FILING DATE: 04-OCT-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: CSV-005.01
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 766 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 -08-539-205A-4

Query Match 38.7%; Score 1564; DB 3; Length 766;
 Best Local Similarity 41.1%; Pred. No. 7.9e-138;
 Matches 339; Conservative 121; Mismatches 232; Indels 132; Gaps 18;
 1 MSPGRRNGPVKRLTVLCAKLVKDFRPPFAKVVWDGSGQCHSTDTVKNITLDPK 60
 1 MNSAQSR----RIRVTIADGLYKRDVFRFPFAVLTVDGE-QTHTTTAKITLNPY 55
 61 WNOHYDLYICKSDSVTISVNNHKKHKKGAGFLGCVRLLSNAINRLKDTGYQL--DLC 118
 56 WNTFEVNDNSTIAIQVFDOKF-KKGGQGLVNLRVGDVLDLAIGDDEMLTRDX 114
 119 KLGPNDNDTVRGQIVVSLQ----- 137
 115 K--SNENTVVGKIIINLSTTAQSTIQVPSSAAGARTQRTSITNDPOSSKSSVSBNPA 172
 138 -----SDRIGTGQVDCSERLDN-----DLPDGEERRRTASGRIOVLNHTTTQ 184
 173 SSRAGSPTRDAPAAASPASSEPRRTSSPQDQYGLRPPGWERRTDNLGRTYVVDNTRSTT 232
 185 WERTP-----RPASYSYSGFRPLSCFVDENPTISGTNGATCGSSDPLAERVRQRH 238
 233 WIRPNLSSVAGAAAEHLHSAS--SANVTEGVQPSSNAA-----RTEASVLV--- 279
 239 RNTMSTLHTPPDPEGYEQRTTQGGQYVFLHTGTGVSTWHDPRVPRDLNIN----- 292
 280 ----SNATTAGSGELPPGWEQYTPGWRPYVDHNTRTTTWDPRRQYRSYGGPNAT 335
 293 -----CEELGELPPGWEIENATGVVVDHNNRTQTDPRLSANLHLVLRNQQLKDQ 347
 336 IQQPVSLQGLPLPSGWEMLTNTARVYVDENKTTTWDPRPSSL-----DQ 384

Query Match 38.7%; Score 1564; DB 4; Length 766;
 Best Local Similarity 41.1%; Pred. No. 7.9e-138;
 Matches 339; Conservative 121; Mismatches 232; Indels 132; Gaps 18;
 1 MSPGRRNGPVKRLTVLCAKLVKDFRPPFAKVVWDGSGQCHSTDTVKNITLDPK 60
 1 MNSAQSR----RIRVTIADGLYKRDVFRFPFAVLTVDGE-QTHTTTAKITLNPY 55
 61 WNOHYDLYICKSDSVTISVNNHKKHKKGAGFLGCVRLLSNAINRLKDTGYQL--DLC 118
 56 WNTFEVNDNSTIAIQVFDOKF-KKGGQGLVNLRVGDVLDLAIGDDEMLTRDX 114
 119 KLGPNDNDTVRGQIVVSLQ----- 137
 115 K--SNENTVVGKIIINLSTTAQSTIQVPSSAAGARTQRTSITNDPOSSKSSVSBNPA 172
 138 -----SDRIGTGQVDCSERLDN-----DLPDGEERRRTASGRIOVLNHTTTQ 184
 173 SSRAGSPTRDAPAAASPASSEPRRTSSPQDQYGLRPPGWERRTDNLGRTYVVDNTRSTT 232
 185 WERTP-----RPASYSYSGFRPLSCFVDENPTISGTNGATCGSSDPLAERVRQRH 238
 233 WIRPNLSSVAGAAAEHLHSAS--SANVTEGVQPSSNAA-----RTEASVLV--- 279
 239 RNTMSTLHTPPDPEGYEQRTTQGGQYVFLHTGTGVSTWHDPRVPRDLNIN----- 292
 280 ----SNATTAGSGELPPGWEQYTPGWRPYVDHNTRTTTWDPRRQYRSYGGPNAT 335
 293 -----CEELGELPPGWEIENATGVVVDHNNRTQTDPRLSANLHLVLRNQQLKDQ 347
 336 IQQPVSLQGLPLPSGWEMLTNTARVYVDENKTTTWDPRPSSL-----DQ 384

Query Match 38.7%; Score 1564; DB 4; Length 766;

380 RQELSQOQOAGHCHRIEVSREIFESYOVVMKORPKDLWK-RLMIKFRGEGLDYGGVA 438
557 RRLLKKQNDIPNFKEMKLRATVLEDSYRIMGVKRAEDFKARLWTEFDGKGLDYGGVA 616
439 REWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMGMAVFG 497
617 REWFLLSKEMFPPYGLFYSATDNYTLQINPNSGLCNEDHLSYFKFGRVAGNAVTHG 676
498 HYDGGTLPFYKOLLGKSTLDDMELVDPDLHNSLVWILENDITGVLDHTFCVEHNAYG 557
677 KLLDGFPIRPFYKMLHKPITLHDMESVDSEYNSLRWILENDPT-ELDLRIIDEELFG 735
558 EIIQHELKPKNGKSTPVNEENKKEYVRLVYNNWRFLRGIEAFLALOKGNEVIPPQHLKTF 617
736 QTHOHELKNGSIVTNNKKKYYILVQWRFNRIQKMAFKEGFFELIPQDLIKIF 795
618 DEKELELIICGLGKIDVNDKVNRLKH-CTPDSNIVKFWKAVEFFDEERARLLQFVT 676
796 DENELELLMGLGDDVDVNDWRETKYKNGYSANHQVQWFKAVLMDSEKIRLLQFVT 855
677 GSSRVPLQGGFKALOGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPYESYEKLYEKL 736
856 GTSRVPNGFAELYGNSGPOSFTVEQW-GTPEKLPRAHTCFNRDLPLPYESPEELWDLKQ 914
737 TAIBETCGF 745
915 MAIENTQGF 923

RESULT 6

US-09-392-205A-6

Sequence 6, Application US/08539205A

Patent No. 6001619

GENERAL INFORMATION:

APPLICANT: Beach, David H.

APPLICANT: Caligiuri, Maureen

TITLE OF INVENTION: Ubiquitin Ligases

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/539,205A

FILING DATE: 04-OCT-1995

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 834 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-392-205A-6

Query Match

Best Local Similarity 33.5%; Score 1353; DB 3; Length 834;

Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPDGWEERTASGRIQYLNHITRTTQWERPTRPASEYSSPG-----RPLS 203
DB 246 LPSGWEERKADKGRTYVYNNHRTTITWIRPIMQLAEDGASGATNSNNHLEPQIRPRS 305
QY 204 C---FVDENTPISGNGATCGGSDPRLAERVRSQRYNMYSTRHLHTPPDLPEGYEQR 260
DB 306 LGSPTVTLAPLEGAKDSPVRAVXKDTLSNPQSPQSPYNSPKQHKVQTQFLPPGEMR 365
QY 261 TTQOGGVYFLHTQTGVSTWHDPRVP-----RDLSNINCEELGPLPPGWEIRNTATGRVYF 315
DB 366 IAPNGRPFIDENTXTTITWEDPRLAKFPVHMSKSLNENDLGPLPPGWEIRHLDGRTFY 425
QY 316 VDHNNRTTQFTDPRLSANLHLVLNQNLKQOQQOQVSVLCPDDTECLTVPR--YKRDLV 373
DB 426 IDHNSKITQWEDPRL-----QN-----PAITGPAVYSREFK 457
QY 374 QKLKILRQELSQQOQAGHCHRIEVSREIFESYOVVMKORPKDLWK-RLMIKFRGEG 432
DB 458 QKYDYFRKLLKXKADIPNRFKMLHNNIPESYRRIMSXKRPDVLKARLWIEFESEKGL 517
QY 433 DYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMG 491
DB 518 DYGGVAREWFFLLSKEMFPPYGLFYSATDNYTLQINPNSGLCNEDHLSYFTTIGRVAG 577
QY 492 MAVFHGHYIDGGFTLPFYKOLLGKSTLDDMELVDPDLHNSLVWILENDITGVLDHTFCV 551
DB 578 LAVFHGKLLDGFPIRPFYKMLHKPITLHDMESVDSEYNSLRWILENDPT-ELDLMF 636
QY 552 EHNAYGBEIIQHELKPKNGKSTPVNEENKKEYVRLVYNNWRFLRGIEAFLALOKGNEVIPPQ 611
DB 637 DEENFGQTYQVDLKPNGSEIMVTNENKEYIDLVIQWRFVNRVQKQNAFLEGFTELLPI 696
QY 612 HLLKTFDEKELELIICGLGKIDVNDKVNRLKH-CTPDSNIVKFWKAVEFFDEERAR 669
DB 697 DLTKIFDENELELLMGLGDDVDVNDWROHSYKNGYC-PNHPVIQWFKAVLLMDAEKRI 755
QY 670 RLLOQVTVSSRVPLQGGFKALOGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPYESYE 729
DB 756 RLLOQVTVTSRVPMNGFAELYGNSGPOSFTVEQWS-PEKLPRAHTCFNRDLPLPYETFE 814
QY 730 KLYEKLTAIBETCGF 745
DB 815 DLREKLLMAVENAQGF 830

RESULT 7

US-09-392-163A-6

Sequence 6, Application US/09392163A

Patent No. 6503742

GENERAL INFORMATION:

APPLICANT: Beach, David H.

APPLICANT: Caligiuri, Maureen

APPLICANT: Nefsky, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205

FILING DATE: 33.5%; Score 1353; DB 4; Length 834;
ATTORNEY/AGENT INFORMATION: 44.8%; Pred. No. 6e-118;
NAME: Vincent, Matthew P. 93; Mismatches 188; Indels 60; Gaps 12;
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
-09-392-163A-6

Query Match 33.1%; Score 1336; DB 2; Length 854;
Best Local Similarity 36.6%; Pred. No. 2.5e-116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;

159 LPDWEERTASGRICLYNHITRTTQWERPTPASEYSSPG-----RPLS 203
160 LPDWEERTASGRICLYNHITRTTQWERPTPASEYSSPG-----RPLS 203
246 LPSGWEERKADKRTYVYNNHRTTWTPIQLAEDGASGATNSNNHLEIPOIRPRS 305
204 C----FVDENTPIGTCGAGCGSSDPLAERVRQRHNTMSRTHLTPDLPPEGHEQR 260
306 LSSPTVTLGAPLEGAKDSPVRAVAKDTLSNPQSPSPYNSFKPKQKVTQSFLLPPGEMR 365
261 TQCGQVYFHTQTGVSTWHDPRVP-----RDLNINCEELGPIPPGWEIRNTATGRVP 315
366 IAPNGRPFIDHNTKTWEDPRKLPVHMRSKTSINPDJGPIPPGWEIRHLDGRTFY 425
316 VDHNNRTTQTPRLSANHLVNLQNQLKQDQVSLCPDDECTVPR--YKRDV 373
426 IDENSKITQWEDPRL-----QN-----PAITGPAVYSREFK 457
374 QKLKILRQELSOQPOAGHCRIEVSRREEIFEESYQVMKMRPKDLWK-RLMIKRGEG 432
458 QKYDYFKKLPKPADIPNPFEMKLRHNNIFEESYRIRMSVRKPDVLRKARLWIESEKGL 517
433 DYGGVAREWLYLLSHEMLNPPYGLFOYSRDDTYTLQINPDSAV-NPEHLSYFHFVGRIMG 491
518 DYGGVAREWFFLLSKEMFNPPYGLFEYSATDNTYTLQINPNSGLCNEDHLSYTFIGRVAG 577
492 MAYFHGYIDGGFTLPFYKQLGKSTLDDMLVDPLDLSNLWILENDITGVLDHFCV 551
578 LAVFHGLDGGFIRPFYKMLGKQTLNDMESVDSEYNSLWILENDPT-ELDLNFCI 636
552 EHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLVYNNRFLRGIEAQFLALOKGFNEVIPQ 611
637 DEENFGQTYQVLDKPNGSEIMVTNENKREYIDLVIQWRVNRVQKQMAFLEGFTLLPI 696
612 HLAKTDEKELELIICGLKIDVNDKWNTRLK--HCTPDNSNIVKWFKAVEFDEERRA 659
697 DLTKIDENELELLMGLGDVDVNDWRQHSIYKNGYC-PNHFVQFWKAVLLMDAKRI 755
670 RLQFVYTGSSRVPLOQFKAQAGAPRTIHOIDACTNNLPKATCNRIDIPYESYE 729
756 RLQFVYTGSRVPMNGFAELYSGNSGQLETTIQWGS-PEKLPRAHTCENRLDLPYETFE 814
730 KLYEKLITAIETTCGP 745
815 DLREKLMAVENAQGF 830

SULT 8

-09-070-060-4
Sequence 4, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Huetad, Carolyn M.
APPLICANT: Ghildyal, Namit
TITLE OF INVENTION: Human E3 Ubiquitin Protein

TITLE OF INVENTION: Ligase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENECA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850-5437
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,060
FILING DATE: 30-APR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/073,839
FILING DATE: 05-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Higgins, Patrick H.
REGISTRATION NUMBER: 39,709
REFERENCE/DOCKET NUMBER: PHM.70312
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302.886.4889
TELEFAX: 302.886.8221
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 854 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-070-060-4

Query Match 33.1%; Score 1336; DB 2; Length 854;
Best Local Similarity 36.6%; Pred. No. 2.5e-116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;

QY 13 KRLRTVLCAK-NLVKKOFFERLPDPFAKVVVGGQCHSTDTVNTLDPKWNQHYDIYIGK 71
DB 9 QCIIVISAKLKENKKWFG-PSPIEVTVD--QSKKTEKCNNTSPKWKQPLTVIVIP 65
QY 72 SDSVTISVMNHKKIHK--QAGFLGCVRLLSNAINRLKDTGYQRLDCKLGPNDNDTVR 129
DB 66 TSKLCFRVWSHQLKSDVLLGLTAGLDIYETLKSNNMKLEBV--VMTQLVGDKEPTETM 122
QY 130 GOIVV---SLQSRDRITGGQV-----DCSRLFDNDLPD-----GWEE 165
DB 123 GDLVCLDGLQVEAEVVTNGETSCSBSTTQDQCRTRDTRVSTNGSEDPVAASGENK 182
QY 166 RRTASGRITQYLNHITRTTQWERPTPASEYSSSPGRLPSCFVDENT-----PIS 213
DB 183 RANGNSPFLSNGGFKSPRPSPRPPP--PTPRPASVNGSPSTNSDSDSGSTGSLPPT 240
QY 214 GTN-----CATCG-----QSSDPLAERVRVSRQHRNYM-----242
DB 241 NTNVTSTSEGTSGLIPLITISGGSGRPLNTVTSQAPLPFGWEQRV-DQHGRVYVYDVH 299
QY 243 -SRTHLHTPPDLPEGYEQRTTQCGVYFHTQTGVSTWHDPRVP-----285
DB 300 EKRTWDRPEPLPFGWERRVDNMGRIVYVDHFTTITWQRTLESVNYEQWLQORSQ 359
QY 286 -----RDL-----SNINCEELGFLPGWEIRNTATGRVYFVDHNNRTTQFTD 327
DB 360 GMAQOFNQRFYIGNQDLFATSONKEFDFLGPLPGWEKRTSDSGRVTYFVNNHTRITQWED 419
QY 328 PRLSANHLVNLQNQLKQDQVSLCP--DDTECLTV---PR-----Y 368
DB 420 PRSQGQ-----LNEKPLPEGWEMRFTVDIGIFVYDHNRRATTYIDPRTGKSALDNGPQIAY 475

369 KRLVQKILR---QELSQOQOAGCHRIEVSREIIEESYRQVMKRPKDLWRLMIK 425
476 VRDFKAKVOYFRWCOQLAMPQ---HIKITVTRKTLFEDSFQOIMSFSPQDLRRLRWI 531
426 FRGEEGLDYGVARWELLYLSHEMLNPNYGLFOYSRDDIYTLQINPDASAVNPEHLSYFHF 485
532 FRGEEGLDYGVARWELLYLSHEMLNPNYGLFOYSRDDIYTLQINPDASAVNPEHLSYFHF 591
486 VGRINGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILENDITGV- 544
592 IGRFIAMALFHGKFDITGFSLPFYKRIILNPKVGLKDLSDIDPEFYNSLIWYKNNIEECG 651
545 LDHTFCVHEHNAIGEIIQHELKENGKSIIPVNEENKKEYVRLVYVNRFLRGITBAOFLALQKG 604
652 LEMYFSDVKELIIGKSHDLKENGNGNLTVEENKKEYIRMAEWRLSRGVEEQOAFPEG 711
605 FNEVTPQHLKTLFDEKELELIICGLKIDVNDWKVNTLKHCTPDSNIVKWFKAVEFFD 664
712 FNEILPQOYLQYFDAKELEVLICGMQEIIDLNDQORHAIYRHYTRTSKQIMWFQVKEID 771
665 EERRARLLQFVGTSSRVLPQGFALQGAAGPRLFTIHOIDACTNNLPKATCENRIDIPP 724
772 NEKMRLLQFVGTGRLPFGVGFADLMGNSGPKFCIEKVGK-ENWLPRSHTCFNRDLDP 830
725 YESYEKLYEKLTAIETTCGFAVE 748
831 YKSYEQLEKLLFAIETEGGQE 854

SULT 9
-09-357-746-4
Sequence 4, Application US/09357746
Patent No. 6087122
GENERAL INFORMATION:
APPLICANT: ZENECA Limited
TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
FILE REFERENCE: PHM-70312.NI
CURRENT APPLICATION NUMBER: US/09/357,746
EARLIER FILING DATE: 1999-07-21
EARLIER FILING DATE: 1998-02-05
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 854
TYPE: PR1
ORGANISM: Mus musculus
-09-357-746-4

Query Match 33.1%; Score 1336; DB 3; Length 854;
Best Local Similarity 36.6%; Pred. No. 2.5e-116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;
13 KRLTVLCAK-NLVKKDPRFLDPPEAKVVDGSGGCHSTDFVKTLDPKXNOHVDLYIGK 71
9 QLGITVISAKLENKKNWFG-PSPEVIVVD--GSKKTEKKNNTNSPKWQPIVTVTP 65
72 SDSVTISVNNKHKHKK--QGAGFLGCVRLLSNAINRLKDTGYQRDLCKLGPNDNDTVR 129
66 TSKLCFRVNSHQTLKSDVLLGTAGLDIYETLKNMKNLEEV--VMTQLVQGDKEPTETM 122
130 GQIVV--SIQSDRIGTGGVV-----DCSLRFDNDLPD-----GWEE 165
123 GDSVCLDGLQVEAEVINGETSCSESTQNDGCRTRDDTRVSTNGSEDPVAASGENK 182
166 RPTASGRIOYLNHTRITQWERPTFPASEYSPGRPLSCFVDENT-----PIS 213
183 RANGNNSPSLNGGPKSPRPPRPP--PTPREPVSNGSPSTNSDGSSTGSLPPT 240
214 GTN-----GATCG-----QSSDPLAERVRSQHRNYM----- 242

Db 241 NTNVTSTSEAGTGLIPLTISGSGRPLNTVSQAPLPPGWEQRY-DQHRGVYVDHV 299
QY 243 -SRTELHTPPLPEGYEORTTQOQVYFLHTQTGVSTWHDPRVP----- 285
Db 300 EKRTTWRPEPLPGWERRVDNNGRIYYVDHFRTTTTQORPTLESVRNYEQWLQSRQLQ 359
QY 286 -----RDL-----SNINCBELGPLPGWEIRNTATGRVYFVDHNNRTTQFTD 327
Db 360 GAMOQFNQRFYGNQDLFATSONKEFDPLGFLPGWEKRTDSNGRYVYFVNHNTRITQWED 419
QY 328 PRLSANLHLNRRNQKLDQOQOQWLSLCP---DDTECLTV---PR-----Y 368
Db 420 PRSQOQ-----LNEKPLEGEMWRTVDIGIFYVDHNRATTYIDPRTGKSAIDNGQIAY 475
QY 369 KRLVQKILR---QELSQOQOAGCHRIEVSREIIEESYRQVMKRPKDLWRLMIK 425
Db 476 VRDFKAKVOYFRWCOQLAMPQ---HIKITVTRKTLFEDSFQOIMSFSPQDLRRLRWI 531
QY 426 FRGEEGLDYGVARWELLYLSHEMLNPNYGLFOYSRDDIYTLQINPDASAVNPEHLSYFHF 485
Db 532 FRGEEGLDYGVARWELLYLSHEMLNPNYGLFOYSRDDIYTLQINPDASAVNPEHLSYFHF 591
QY 486 VGRINGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILENDITGV- 544
Db 592 IGRFIAMALFHGKFDITGFSLPFYKRIILNPKVGLKDLSDIDPEFYNSLIWYKNNIEECG 651
QY 545 LDHTFCVHEHNAIGEIIQHELKENGKSIIPVNEENKKEYVRLVYVNRFLRGITBAOFLALQKG 604
Db 652 LEMYFSDVKELIIGKSHDLKENGNGNLTVEENKKEYIRMAEWRLSRGVEEQOAFPEG 711
QY 605 FNEVTPQHLKTLFDEKELELIICGLKIDVNDWKVNTLKHCTPDSNIVKWFKAVEFFD 664
Db 712 FNEILPQOYLQYFDAKELEVLICGMQEIIDLNDQORHAIYRHYTRTSKQIMWFQVKEID 771
QY 665 EERRARLLQFVGTSSRVLPQGFALQGAAGPRLFTIHOIDACTNNLPKATCENRIDIPP 724
Db 772 NEKMRLLQFVGTGRLPFGVGFADLMGNSGPKFCIEKVGK-ENWLPRSHTCFNRDLDP 830
QY 725 YESYEKLYEKLTAIETTCGFAVE 748
Db 831 YKSYEQLEKLLFAIETEGGQE 854

RESULT 10
US-09-070-060-3
Sequence 3, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Hustad, Carolyn M.
APPLICANT: Grindval, Namit
TITLE OF INVENTION: Human E3 Ubiquitin Protein
TITLE OF INVENTION: Ligase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENECA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850-5437
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,060
FILING DATE: 30-APR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/073,839
FILING DATE: 05-FEB-1998

458 PRTGK---ALDNGPQI-----AYRDFKAKVQFRFCWQOLA 492
385 QQPOAGHCHIEVSREEIPEESYRQWVKRDKLWRLMIKFRGEGLDYGVAREWLYL 444
493 MPQ---HIKITVTRKTLFEDSPQIMSFSPQDLRRRLWVIFPGEGLDYGVAREWFL 548
445 LSHEMLNPFYGLFQYGRDDIYTIQINPDSAVNPEHLSYFHFVGRIMGMVAFHGHVIGGF 504
549 LSHEVLNPFYCLFEYACKONYCLOINPASYINPDHLKYFRFRIGRIAMALFHGKIDTGF 608
505 TLFPYKQLKSTLDDMELVDPDLNSLVILENDITGV-LDHTFCVHNAYGEIIOHE 563
609 SLFPYKRLNPKVGLDLSIDPEFYNSLIWKENNIECDLEMYFSVDEKELGKSHD 668
564 LKNGKSIPIVNEENKKEYRVLRYNNVREFLAGIEAFLQKGENEVIPOHLLKTDKELE 623
669 LKNGGNILVTENKEEYIRMAVRLSGVSEQTAFFEGNEILPOOYLOYFDAKELE 728
624 LIICGLKIDVNDWKNYTRKHCTPDSPNIVKFWKAVEFFDEEREARLLQFVTGSSRVPL 683
729 VLLCGMQEIDLNDQWRHATYRHVARTSKQIMFWQFVKEIDNEKEMRLIQFVTGCRPLV 788
684 QGPKALQGAAGPRLFIHQIDACTNNLPKATCFNRIDIPPEYSEKLYEKLTAIBETC 743
789 GGFADLMSNGPQKFCIEKVGK-ENWLPRESHTCFNRDLDPYKSYEQKELLLFAIBETE 847
744 GFAVE 748
848 GFQGE 852

SULT 12

-08-630-916A-48
Sequence 48, Application US/08630916A
Patent No. 6011137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide

-08-630-916A-48

Query Match 32.2%; Score 1298.5; DB 3; Length 906;
Best Local Similarity 42.0%; Pred. No. 9e-113;
Matches 271; Conservative 101; Mismatches 216; Indels 57; Gaps 9;
QY 122 RNDNDTVRGQIVVLSQSDRDTGTGGQVVDSCRLFDNDLPDQWEERRTASGRIOYLNHTR 181
DB 301 PPNNTTSLPAPATPAEGEESTSGTQPLPAAQAPDALPAGWEQRELPNGRVVYVDHNTK 360
QY 182 TTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGQSSDPRLAERRVRSQRHN- 240
DB 361 TTTWERPLPGWKEKTDPRGRFYVDHNT-----RITWQRPATAEYVRYVYEQWQSRNQ 414
QY 241 -----YMSRTHLHP-DLPPEGYQORTQQQVYFLHTQTGVSFWHDPV 284
DB 415 LOGAMQHSQRFYQVWSASTDHPGLPFGWEKR-QDNGRVVYVNHNTRTTQWEDPT 473
QY 285 PRDLNINCEELGPPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQNL 344
DB 474 QGMIOE-----PALPPQWEMKYTSEGVRVFDHNTTITTFKDPGFE----- 516
QY 345 KDQOQQQVVSCLPDDTECLTVRYKRDVLQKLTILRQELSQQQPAQGHCHIEVSREIFE 404
DB 517 -----SGTKQSGPAYDRSFRWKYHQR-FLCHSNALPSHWKISVSQRTLFE 562
QY 405 ESYROVMKVRPKDLWKRLMIKFRGEGLDYGVAREWLYLISHMLNPPYGLFQYSRDDI 464
DB 563 DSFQIMMVKYDILRRRLYIIMRGEGLDYGVAREWFFLLSHEVLNPNMYCLPEYAGKN 622
QY 465 YTLQINPDSAVNPEHLSYFHFVGRIMGMVAFHGHVIGDGTFFPYKQLKSTLDDMEL 524
DB 623 YCLQINPASSINPDHLTYFRFGRIFAMALYHGFIDTGTFTLPPYKEMLNKRTPLKDL 682
QY 525 VDPDLHNSLVILENDITGV-LDHTFCVHNAYGEIIOHELKHGKSIPIVNEENKKEYR 583
DB 683 IDPEFYNSIVMIKENNLEECLELYFQDMELGKVTTHLKEGGSIRVTEENKEEYIM 742
QY 584 LYVNRFLRGIEAQLALQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDWKNYTR 643
DB 743 LLTDWRFTRGVEEQTKAFLDGNEVAPLEWLRVDFEKELEMLCGMQEIDMSDWQKSTIY 802
QY 644 KHCPTDSNIVKFWKAVEFFDEEREARLLQFVTGSSRVPLQGFALQGAAGPRLFIHQI 703
DB 803 RHYTKNSQIQWFWQVVKEMDNKRIELLOFVTGTCRLPVGFGFAELIGSNGPQKFCIDKV 852
QY 704 DACTNNLPKATCFNRIDIPPEYSEKLYEKLTAIBETCQFAVE 748
DB 853 GKET-WLPRESHTCFNRDLDPYKSYEQKELLYAIBETGFGQE 906

RESULT 13

US-08-630-916A-46
Sequence 46, Application US/08630916A
Patent No. 6011137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide

SOFTWARE: Patent In Release #1.0, Version #1.30

AFFILIANT: ECKSTEIN, JENS W
APPLICANT: Draetta, Giulio

Db 626 CILDVHFPMVYRKL MGKGLFVDLGDSHPVLYQSLKDLLEYVGNVEDDMMITFQISQTN 685

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

% protein - protein search, using sw model

on on: February 20, 2004, 15:27:02 ; Search time 33.0523 Seconds
(without alignments)
4738.482 Million cell updates/sec

itle: US-10-009-945-4

arfact score: 4038

sequence: 1 MSNPGRRNGPVKRLTVLC.....EKLYEKLTAETTCGFAVE 748

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

sarched: 801455 seqs, 209382283 residues

otal number of hits satisfying chosen parameters: 801455

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	4031	99.8	748	12	US-10-021-660-81
2	3939.5	97.6	735	12	US-10-313-955-2
3	3001	74.3	722	15	US-10-097-534-14
4	2775	68.7	514	12	US-09-764-875-819
5	1564	38.7	766	12	US-10-313-955-4
6	1530.5	37.9	832	12	US-10-032-585-7296
7	1453.5	36.0	869	15	US-10-128-714-8162
8	1370.5	33.9	911	15	US-10-205-823-279
9	1355.5	33.6	725	15	US-10-185-050-126
10	1354.5	33.5	854	15	US-10-205-823-277
11	1354.5	33.5	927	15	US-10-097-534-15
12	1353.5	33.5	995	15	US-10-097-534-9
13	1353.5	33.5	995	15	US-10-205-823-275
14	1353	33.5	834	12	US-10-313-955-6
15	1340.5	33.2	759	15	US-10-128-714-3162

16	1338.5	33.1	854	12	US-10-287-218-3
17	1301	32.2	739	12	US-10-374-979-89
18	1301	32.2	739	15	US-10-097-534-10
19	1301	32.2	739	15	US-09-919-039-235
20	1298.5	32.2	870	15	US-10-097-534-12
21	1298.5	32.2	906	15	US-10-185-050-48
22	1209	29.9	474	11	US-09-774-639-371
23	1209	29.9	474	11	US-09-968-730-249
24	1134.5	28.1	683	15	US-10-185-050-46
25	1134.5	28.1	684	15	US-10-097-534-11
26	1050	26.0	1562	15	US-10-097-534-13
27	1019.5	25.2	380	15	US-10-307-956-32
28	1008.5	25.0	375	15	US-10-307-956-31
29	869	21.5	733	15	US-10-097-534-16
30	835.5	20.7	1094	15	US-10-043-487-300
31	835.5	20.7	1488	15	US-10-043-487-285
32	775.5	19.2	277	10	US-09-925-300-1628
33	733	18.2	358	15	US-10-268-036-5
34	682	16.9	276	10	US-09-925-300-1527
35	665	16.5	125	9	US-09-764-870-384
36	665	16.5	125	12	US-09-764-875-1110
37	665	16.5	125	15	US-10-125-540-384
38	499	12.4	823	15	US-10-097-534-25
39	445.5	11.0	1050	15	US-10-097-534-24
40	445.5	11.0	1050	15	US-10-097-534-28
41	445.5	11.0	1054	15	US-10-097-534-29
42	422	10.5	1024	15	US-10-097-534-30
43	407.5	10.1	4861	10	US-09-919-497-70
44	407.5	10.1	4861	15	US-10-097-534-26
45	407.5	10.1	4861	15	US-10-146-473-49

ALIGNMENTS

RESULT 1
US-10-021-660-81
; Sequence 81, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 81
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-81

Query Match 99.8%; Score 4031; DB 12; Length 748;
Best Local Similarity 99.9%; Pred No. 0;
Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSNPGRRNGPVKRLTVLCANLVKDDFFRLPDPFAKVVDGSGQCHSTDTVKTLDPK 60

Db 1 MSNPGRRNGPVKRLTVLCANLVKDDFFRLPDPFAKVVDGSGQCHSTDTVKTLDPK 60

Qy 61 WQHYDLYIGKSDSVTISWNHKKIHKQAGAGCGVLLSNAINRLKDTGYORLDLCKL 120

Db 61 WQHYDLYIGKSDSVTISWNHKKIHKQAGAGCGVLLSNAINRLKDTGYORLDLCKL 120

121 GPNNDIVRGQIVVLSQSRDRIGTGQGVVDCGRLPDNDLPDQWEERRTASGRIOYLNHIT 180
 121 GPNNDIVRGQIVVLSQSRDRIGTGQGVVDCGRLPDNDLPDQWEERRTASGRIOYLNHIT 180
 181 RTQWERPTPASEYSPGRLSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 240
 181 RTQWERPTPASEYSPGRLSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 240
 241 YMSRTHLTPDLPDPEGVEQRTTQGGVFLHQTGVSTWHDPRVDRDLSNINCELGPLP 300
 241 YMSRTHLTPDLPDPEGVEQRTTQGGVFLHQTGVSTWHDPRVDRDLSNINCELGPLP 300
 301 PGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLQDQOQQVSLCPDDT 360
 301 PGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLQDQOQQVSLCPDDT 360
 361 ECLTVPRYKRDVLQKILRQELSLQOQPGAGHCRIFVSREEIFEESYRQVMKRPDLWK 420
 361 ECLTVPRYKRDVLQKILRQELSLQOQPGAGHCRIFVSREEIFEESYRQVMKRPDLWK 420
 421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 480
 421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 480
 481 SYFHFVGRIMGMAVFGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLWILEND 540
 481 SYFHFVGRIMGMAVFGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLWILEND 540
 541 ITGVLDTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLVNNRFLRGIEAQFLA 600
 541 ITGVLDTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLVNNRFLRGIEAQFLA 600
 601 LQKGFNEVIPHLLKTDFEKELELIICGLGKIDVNDKVNTRPKHCTPDSNIVKWFKAV 660
 601 LQKGFNEVIPHLLKTDFEKELELIICGLGKIDVNDKVNTRPKHCTPDSNIVKWFKAV 660
 661 EFPDEERARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 720
 661 EFPDEERARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 720
 721 DIPPYSEYKLYEKLTAIEETCGFAVE 748
 721 DIPPYSEYKLYEKLTAIEETCGFAVE 748

SULT 2
 -10-313-955-2
 Sequence 2, Application US/10313955
 Publication No. US20030199036A1
 GENERAL INFORMATION:
 APPLICANT: Beach, David H.
 Nefsky, Bradley
 Caligiuri, Maureen
 TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/313,955
 FILING DATE: 05-Dec-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/392,163

FILING DATE: <Unknown>
 APPLICATION NUMBER: US 08/539,205
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: CSV-005.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 735 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-313-955-2
 Query Match 97.6%; Score 3939.5; DB 12; Length 735;
 Best Local Similarity 98.0%; Pred. No. 0;
 Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;
 QY 1 MSNPGRRNGPVKRLTLVLCANLVKDDFFRLPDPFAKVVVDGSGGCHSTDTVKNLTDPK 60
 DB 1 MSNPGRRNGPVKRLT-----GLPDPFAKVVVDGSGGCHSTDTVKNLTDPK 47
 QY 61 WNHQYDLVIGKSDSVTISVNNKXIHKKQAGFLGCVRLLSNAINRLKDTGYQRLDCKL 120
 DB 48 WNHQYDLVIGKSDSVTISVNNKXIHKKQAGFLGCVRLLSNAINRLKDTGYQRLDCKL 107
 QY 121 GENDNDTVRGQIVVLSQSRDRIGTGQGVVDCSRLPDNDLPDQWEERRTASGRIOYLNHIT 180
 DB 108 GENDNDTVRGQIVVLSQSRDRIGTGQGVVDCSRLPDNDLPDQWEERRTASGRIOYLNHIT 167
 QY 181 RTQWERPTPASEYSPGRLSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 240
 DB 168 RTQWERPTPASEYSPGRLSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 227
 QY 241 YMSRTHLTPDLPDPEGVEQRTTQGGVFLHQTGVSTWHDPRVDRDLSNINCELGPLP 300
 DB 228 YMSRTHLTPDLPDPEGVEQRTTQGGVFLHQTGVSTWHDPRVDRDLSNINCELGPLP 287
 QY 301 PGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLQDQOQQVSLCPDDT 360
 DB 289 PGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLQDQOQQVSLCPDDT 347
 QY 361 ECLTVPRYKRDVLQKILRQELSLQOQPGAGHCRIFVSREEIFEESYRQVMKRPDLWK 420
 DB 348 ECLTVPRYKRDVLQKILRQELSLQOQPGAGHCRIFVSREEIFEESYRQVMKRPDLWK 407
 QY 421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 480
 DB 408 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 467
 QY 481 SYFHFVGRIMGMAVFGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLWILEND 540
 DB 468 SYFHFVGRIMGMAVFGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLWILEND 527
 QY 541 ITGVLDTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLVNNRFLRGIEAQFLA 600
 DB 528 ITGVLDTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLVNNRFLRGIEAQFLA 587
 QY 601 LQKGFNEVIPHLLKTDFEKELELIICGLGKIDVNDKVNTRPKHCTPDSNIVKWFKAV 660
 DB 588 LQKGFNEVIPHLLKTDFEKELELIICGLGKIDVNDKVNTRPKHCTPDSNIVKWFKAV 647
 QY 661 EFPDEERARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 720
 DB 648 EFPDEERARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 707
 QY 721 DIPPYSEYKLYEKLTAIEETCGFAVE 748

708 DIPPYESYEKLEKLLTAIBETCGFAVE 735

RESULT 3

3-10-097-534-14

Sequence 14, Application US/10097534

Publication No. US20030049607A1

GENERAL INFORMATION:

APPLICANT: GREENER, TSVIKA

APPLICANT: MOSKOWITZ, HAIM

APPLICANT: REISS, YUVAL

APPLICANT: ALROY, IRIS

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL

FILE REFERENCE: PLV-001.01

CURRENT APPLICATION NUMBER: US/10/097,534

PRIOR FILING DATE: 2002-03-12

PRIOR FILING DATE: 2002-03-12

PRIOR FILING DATE: 2001-03-12

PRIOR FILING DATE: 2001-07-31

PRIOR FILING DATE: 2001-12-07

NUMBER OF SEQ ID NOS: 71

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 14

LENGTH: 722

TYPE: PRT

ORGANISM: Homo sapiens

-10-097-534-14

Query Match

Best Local Similarity 74.3%; Score 3001; DB 15; Length 722;

Matches 563; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

12 VKRLTVLCAKLVKKDFRPPFAKVVVDGSGCHSDTIVKNTLDPKWOHVDLYGK 71

4 IKRLTVLCAKLVKKDFRPPFAKVVVDGSGCHSDTIVKNTLDPKWOHVDLYGK 63

72 SDSVTSVWNHKKIHKQAGFGLGCVRLLSNAINRLKDTGYORLCLKLPNDNDVRGQ 131

64 TDSITISVWNHKKIHKQAGFGLGCVRLLSNAINRLKDTGYORLCLKLPNDNDVRGQ 123

132 IVVLSQSRDRIGTGQGVDCSLFDNDLPGDWEERRTAGRIQYLNIHTRTQWERTRP 191

124 IVVLSQTRDRIGTGQGVDCRGLLENE-----GTVY----- 154

192 ASYSFGRPLSCFVDENTPIGTINGATCG-----QSSDPLAERRVRSQRHNYM 242

155 --BDSGGRPLSCFMEEPAPYTDSTGAAAGGNCRFVESPQDQRLQRLNPDVRGSL 212

243 ----SRTHLTPDLPEGVEQRTQGGVYFLHTQTGVTWHDPRVPRDLNINCBLGP 298

213 QTQNRPHGHSPELPEGVEQRTTQGGVYFLHTQTGVTWHDPRVPRDLNINCBLGP 272

299 LPPGWEIRNATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQNOLKDDQOQQVV--S 354

273 LPPGWEIRNATGRVYFVDHNNRTTQTDPR-----LHMHMHQQLKEPSQPLPSEGS 328

355 LCPDDTECLTVPRYKEDLVQKILRQELSQQQAGCHRIEVSREEIFEESYQVMKMR 414

329 L-----EDELPAQRYEDLVQKILRQELSLQOQAGCHRIEVSREEIFEESYQVMKMR 385

415 PKDLWRLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSA 474

386 PKDLWRLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSA 445

475 VNPHELSYFHFVGRIMGAVFHGHDGFTLPPYKQLLGSITLDDMELVDPDLHNSLV 534

446 INPDHLSYFHFVGRIMGAVFHGHDGFTLPPYKQLLGSITLDDMELVDPDLHNSLV 505

535 WILENDITGVLDHTFCVEHNAYGEIIQHELKPNKSGIPVNEENKKEYVRLYVNNRFLRGI 594

Db 506 WILENDITGVLDHTFCVEHNAYGEIIQHELKPNKSGIPVNEENKKEYVRLYVNNRFLRGI 565

Qy 595 EAQFLALQKGFNEVFPQHLKTFDEKELELIICGLGKIDVNDWKVNTRLKCHCTPDSNIVK 654

Db 566 EAQFLALQKGFNEVFPQHLKTFDEKELELIICGLGKIDVNDWKVNTRLKCHCTPDSNIVK 625

Qy 655 WFKWAVEFFDEERARLLQFVTSRSSRVPLOGFKALO---GAAGPRLFTIHOIDACTNNLP 711

Db 626 WFKWAVEFFDEERARLLQFVTSRSSRVPLOGFKALO---GAAGPRLFTIHOIDACTNNLP 685

Qy 712 KAHTCFNRIDIPPPYESYEKLEKLLTAIBETCGFAVE 748

Db 686 KAHTCFNRIDIPPPYESYEKLEKLLTAIBETCGFAVE 722

RESULT 4

US-09-764-875-819

; Sequence 819, Application US/09764875

; Publication No. US20040018969A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: P202

; CURRENT APPLICATION NUMBER: US/09/764,875

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1249

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 819

; LENGTH: 514

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-875-819

Query Match 68.7%; Score 2775; DB 12; Length 514;

Best Local Similarity 99.8%; Pred. No. 1.7e-248;

Matches 513; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 235 SQRHNYMSRTHLHTFPDLPEGVEQRTTQGGVYFLHTQTGVTWHDPRVPRDLNINCE 294

Db 1 SQRHNYMSRTHLHTFPDLPEGVEQRTTQGGVYFLHTQTGVTWHDPRVPRDLNINCE 60

Qy 295 ELGPLPPGWEIRNATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQNOLKDDQOQQVVS 354

Db 61 ELGPLPPGWEIRNATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQNOLKDDQOQQVVS 120

Qy 355 LCPDDTECLTVPRYKEDLVQKILRQELSQQQAGCHRIEVSREEIFEESYQVMKMR 414

Db 121 LCPDDTECLTVPRYKEDLVQKILRQELSQQQAGCHRIEVSREEIFEESYQVMKMR 180

Qy 415 PKDLWRLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSA 474

Db 181 PKDLWRLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSA 240

Qy 475 VNPHELSYFHFVGRIMGAVFHGHDGFTLPPYKQLLGSITLDDMELVDPDLHNSLV 534

Db 241 VNPHELSYFHFVGRIMGAVFHGHDGFTLPPYKQLLGSITLDDMELVDPDLHNSLV 300

Qy 535 WILENDITGVLDHTFCVEHNAYGEIIQHELKPNKSGIPVNEENKKEYVRLYVNNRFLRGI 594

Db 301 WILENDITGVLDHTFCVEHNAYGEIIQHELKPNKSGIPVNEENKKEYVRLYVNNRFLRGI 360

Qy 595 EAQFLALQKGFNEVFPQHLKTFDEKELELIICGLGKIDVNDWKVNTRLKCHCTPDSNIVK 654

Db 361 EAQFLALQKGFNEVFPQHLKTFDEKELELIICGLGKIDVNDWKVNTRLKCHCTPDSNIVK 420

Qy 655 WFKWAVEFFDEERARLLQFVTSRSSRVPLOGFKALO---GAAGPRLFTIHOIDACTNNLP 714

Db 421 WFKWAVEFFDEERARLLQFVTSRSSRVPLOGFKALO---GAAGPRLFTIHOIDACTNNLP 480

Qy 715 TCFNRIDIPPPYESYEKLEKLLTAIBETCGFAVE 748

Db 481 TCFNRIDIPPPYESYEKLEKLLTAIBETCGFAVE 514

RESULT 5

3-10-313-955-4

Sequence 4, Application US/10313955

Publication No. US20030199036A1

GENERAL INFORMATION:

APPLICANT: Beach, David H.

Nefsky, Bradley

Calligiuri, Maureen

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/10/313,955

APPLICATION NUMBER: US/10/313,955

FILING DATE: 05-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/539,205

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 766 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

-10-313-955-4

Query Match

Best Local Similarity 41.1%; Score 1564; DB 12; Length 766;

Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;

1 MSNPGRRNGPVKRLTVLCAKLVKKDFRFPDPFAKVVVDGSGQCHSTTVKNTLDPK 60

1 MSNSAQR-----KIRVTIWAADGLYKRDVFRFPDPFAVLTVDGE-QTHTTTAKTLNLY 55

61 WNOHYDLIGKSDSVTISVMNHKKHKQAGFLGCVRLLSNAINRLKDTGVQRL--DLC 118

56 WNETFEVNTDNSTIALQVFDQKF-KKKGQGLGVINLRVGDVLDAIGDEMTRDLK 114

119 KLOPNDNTVRGQIVVLSLQ----- 137

115 K--SNENTVVHGKIIINLSTTAOSTLQVPSSAASGARTORTSITNDPOSSKSSVS RNPA 172

138 -----SRDRIGTGQVQVDCGRLFN-----DLPDGEERTASGRIOYLNHITRTTQ 184

173 SSRAGSPTRDNAPASPASERPTSSFDQYGRLLPGWERTDNLGYVVDHNTASTT 232

185 WERTP-----RPASYSFGPRPLSCFVDENTPISGTNCAIGQSSDPRLAERRVRSQRH 238

233 WIRPNTSSVAGAAAALHSSAS--SANVTEGVQVPSSNAA-----RRTEASVLT--- 279

239 RNYMSRTHLTPDLPGEYEQRTTQGGQVYFLHTGTGVSTWHDPRVPRDLNIN----- 292

Db 280 ----SNATTAGSGELPPGWEQRYTFEGRPYFVDNTRTTTWDPRRQYIRSYGGPNAT 335

QY 293 -----CEELGPLPPGWEIRNTATGRVYFVDHNNRTTQTPRJSANLHLVNRQNKDQ 347

Db 336 IQQQPVSQLPGLPSGWEMLNTARVYFVDHNTKTTTWDPRLPSSL-----DQ 384

QY 348 QQQQVVSCLCFDDTECLTPRYKRDVLQKILKRLRQLSQ--QPQAGHCRLEVSREEIFE 405

Db 385 -----NVPOYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHNIFED 425

QY 406 SYRQVMKRPKDLWKRLMKFRGEEGLDYGAVREWLYLLSHENLNFYGLFQYSRDDIY 465

Db 426 SYAETMRQSATDLKKRLMKFDGEDGLDYGSLREYFLLSHEMFNFYCLFYSSVDNY 485

QY 466 TLQINPDSAVAPRHLSEYFHFVGRIMGMAVFGHVIDGGFTLPFYKQLLGKSITLDDMELV 525

Db 486 TLQINPHSGINPEHLNYPKFIGRVIGLAIFHRFRRVDAFFVVSFYKMLQKKVLQDNESM 545

QY 526 DPDLNSLWILENDITGVLDHTFCVEHNAYGEIIQHELKPNKSKI PVNEENKKEYVRLY 585

Db 546 DAERYSLVWLNDNDITGVLDLTSVEDNCFGEVVTIDLKENGRIEVTENKREYVDLV 605

QY 586 VNWRFLEHIEAQLAQGFNEVTPQHLTKTFDKELELICGLGKIDVNDWKVNTLKH 645

Db 606 TYW-IQKRIEQFNAPHEGFSLEIPOELINVDRELELLIGGISEIDMEDWKKHKDYRS 664

QY 646 CTDPDGNIVKWFKAVEFDEERRARLLQFVTGSSRVLPQGFKALQGAAGPRLFTIHOIDA 705

Db 665 YSENDQLIKFWELMDENSKSRLQLFTGTGSRIPVNGFKDLQSGDGRPKRTIEKAGE 724

QY 706 CTNNLPKHAHTCFNRIDIPPYSYKLYEKLTAIETCGFAVE 748

Db 725 -FNKLPKARTCFNRDLDPPTYSKKDLQKLSIAVEETIGFQGE 766

RESULT 6

US-10-032-585-7296

; Sequence 7296, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; APPLICANT: Howard, Bussey

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 7296

; LENGTH: 832

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-032-585-7296

Query Match

Best Local Similarity 37.9%; Score 1530.5; DB 12; Length 832;

Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

QY 6 RRRNGFVKRLRTVLCAKLVKKDFRFPDPFAKVVVDGSGQCHSTDTVKNTLDPKQNHY 65

Db 9 RSNNTTTINVKWAABESLYKRDVFRQDPDPFAVLTVDGS-QTKTTITAKTLNLYWNEIF 67

QY 66 DLYIGKSDSV-TISWNHKKHKQAGFLGCVR-----LLSNAINRLKDTGVQRLDLC 119

Db 68 N-FOAKEDSILVIQVFDQKF-KKKGQGLGVINLRVGDVIDLSLNSSEETITR--DLXX 123

QY 120 LQPNNDNTVRGQIVVVSQSRDRIGTGQGVVD----- 150

Db 124 --SNENLAVSGKIIWVI-SHNENSGGVVTTATRTGTGSSNNIATITSGVNNLRIGSA 180

151 -----CSRLPND--- 157
181 TTTANSTAASSDATVANGSGPTSSLPPIGQGHPTAATPGGAAGAAAGROQSSPED 240
158 ---DLPQGWERRTASGRIOYLNHITTTTOWERPTPASE----- 194
241 QYGLPLPGWERRTDFORTYVYVHNSTTTTQRPALHQSETERGQORQSETEAERQRHG 300
195 YSSPGR---PLSCFVDENTPISGTNGATCQSS-----DPLAERRVRQRHRN 240
301 RTLPGEQSVSPL-----PTGSGNSITSGNTVTVASGANTFVNPAAVVAASGATTS 352
241 YMSRTHLHTPDLPEGVEQRTTQGGVYFHTGTGVSTWHDPRVPRDLNIN----- 292
353 GLG-----ELPSSGEQRTTEGRPYFVDHNTTITTTWDPDRQQYVIRTPGNTTIOQ 404
293 -CEBLGPLPFGWEIRNTATGRVYFVDHNRHTTQTPRLSANLHLVLRNQKLDQOQQO 351
405 PVSQGLPLPGWENLNTARVYFVDHNTKITTWDDPRLPSSL-----DQ----- 449
352 VVSLCPDTECLTVPRYKRDVLQKILRQELSOQOQ-----AGHCRIEVSREEIFESY 407
450 -----NVPOYKDFPRKVIYFR-----SQPALRILPGQCHIKVREDHIFEDSY 492
408 ROYMKRPEKDLWKRLMKPGEGLDYGVARWLYLLSHEMLNPPYGLFQVSRDDIYTL 467
493 QELNRQTPEDLKKRLMKPGEGLDYGVSREFFFLSHDMFNPFYCLFEYSSSHNTYL 552
468 QINPDSAVNPEHLSYHFVGRIMGMAVFHGYIDGGFTLPFYKQLLKGKSLTLDDELVD 527
553 QINPNSGINPEHLNYPKFGVGLGVFHRRLDFAFFVGLYKWLKVKVVLQMEGYDA 612
528 DLHNSLWILENDITGVLDHFTFCVEHNAIGEIOHELKPKNGKSIPVNEENKYYVRLVYN 587
613 EFRSLKWLINDITGLDITFSAEESFEIYEVLDKPGGRDIEVTENKHEYVELITE 672
588 WRFLRGIEAQLAQGFNEVIVPQHLKTFDEKELELILCGIKGKIDVNDKVNTELKHC 647
673 WRISKRVQEQKAFIDGNELIQELVNVFDERELELLGGLAEIDCEDKWKHTDYRGY 732
648 PDSNIVKFWKAVEFDEERRARLLFOVTGSSRVPLQGFALQGAAGPRLFTIHOIDACT 707
733 ENDQVQWFKCINWDSEQKRLLOFTGTSRIPVNGFKDLQSGDGRPRFTIEKAGE-A 791
708 NNLPKAHTCFNRIDIPYSEYKLEKLATATEECGRAVE 748
792 NQLPKSHTCNFRVLDPPYTDYBSLKOKLTLAVEETVGFQGE 832

SULT 7

-10-128-714-8162
Sequence 8162, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wengqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIORITY APPLICATION NUMBER: US 60/285,697
PRIORITY FILING DATE: 2001-04-23
PRIORITY APPLICATION NUMBER: US 60/287,066
PRIORITY FILING DATE: 2001-04-27
PRIORITY APPLICATION NUMBER: US 60/295,890
PRIORITY FILING DATE: 2001-06-05
PRIORITY APPLICATION NUMBER: US 60/303,899
PRIORITY FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8162
LENGTH: 869
TYPE: PRT
ORGANISM: *Aspergillus fumigatus*
US-10-128-714-8162

Query Match 36.0%; Score 1453.5; DB 15; Length 869;
Best Local Similarity 36.2%; Pred. No. 2,1e-125;
Matches 332; Conservative 114; Mismatches 223; Indels 243; Gaps 21;

QY 11 PVKRL-L-TVLCAKLVKQDFR-----LPDPFAKVVVDGSGQCH 48
Db 15 PASSRLCRVIAADGKYKEDVFRKSYLAILLVRLTSLTSGAFAGPDPFPAVATVGE-QTH 73
QY 49 STDVTNLTDPKXNQHLYIGKSDSVTISVMNHKKIHKQAGFLGCVRL----- 100
Db 74 TTSVIKTLNPNYWNEMFWRVNEDESILAIQFDQKF-KKKQOQFLGVINRIGDVIDLQ 132
QY 101 -----SNAINRLK-----DTGQRLDLCGLPNDNDTVRGQIVVSL----- 136
Db 133 MGGDGBESLPIRSHCDVSRILTFRLLLLETEMLTRDLKK--SNDNLVWVGKLIINLTSLST 190
QY 137 -----QSRDRIGTGGGVDCS----- 152
Db 191 PNTNQANGLHRSHVSSSTSSGLVPQVAFSSSHPAASGTAPVDPASNPSPNLPQRPVSTR 250
QY 153 -----RLFDND-----LPGWEERRTASGRIOYLNHITTTQ 184
Db 251 PSTAAPASAAAGAAVNSHSGRTNLSFEDSQRLPAGWERREDNLGRTYVVDNTRITTT 310
QY 185 WERPTRPASEYSSPCRLPSCFVDENTPISGTNGATCGOSSPRLAERRVRSORHNTM-- 242
Db 311 W--TRPSSNYNEHAQ-----RSQREANMQLERRAHQSRMLP 344
QY 243 -----SRTHLTHPP-----DLPEGYEORTTQOQVY 268
Db 345 EDTRGANSNLPESSQOAHPPAGSANAASVMATGATTAGTGELPPGWEQRTTPEGPY 404
QY 269 FLHTQTGVSTWHDPRVRL-----SNIN-----CEBLGPLPGWEIRNTATGRVY 314
Db 405 FVDHNTTITTTWDPDRQQYIRMYGQNGANGTNTTIOQOQVPSQLGFLPSGWMRLNTARVY 464
QY 315 FVDHNRHTTQTPRLSANLHLVLRNQKLDQOQQOQVSLCPDDTECLTVPRYKRDVLQ 374
Db 465 FVDHNTKITTWDDPRLPSSL-----DQ-----GVQYKDRFR 497
QY 375 KLIKRLRQELSOQOQ-----OAGHCRIEVSREEIFESYRQVQMKRPKDLWKRLMKFRGEE 430
Db 498 KLIYFR-----SQPALRIMSQQCHVKVRNNIFEDSYAETMRQSASDLKKRLMKPDGED 552
QY 431 GLDYGVARWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHLSYHFVGRIM 490
Db 553 GLDYGGLSGEFFFFLLSHEMFNPFYCLFEYSADHNTTITQINPHSGVNPHEHLNPKFGIRV 612
QY 491 GNAVPHGYIDGGFTLPFYKQLLKGKSLTLDDELVDLPDLHNSLWILENDITGVLDHFTFC 550
Db 613 GLAIFHRRLDLSFFTFAGFYKWLKVKVSLQDMEGVDEDLHRLNTWTWNDIEGVLELTF 672
QY 551 VEHNAYGEIIQHELKPKNGKSTPVNEENKKEYVRLVYNVWFLRGIEAQLQGFNEVIP 610
Db 673 VDDEKFGERRTIDLPKGGDRIPVTNENKAEYVRLVTEWKIVKRVVEQFNAFMGFGNELIP 732
QY 611 QHLKTFDEKELELIIICGLKIDVNDKVNTRKLGCTPDSNIVKFWKAVEFDEERRAR 670
Db 733 ADLVNVFDERELELLIGIADIDVDDMKKHTDVRGYQESDEVIONFWKIVRSWDAEQKS 792
QY 671 LLOFTVGTSSRVPLQGFALQGAAGPRLFTIHOIDACTNNLPKHAHTCFNRIDIPYSEYK 730
Db 793 LLQFTTGTSRIPVNGFKDLQSGDGRPRFTIEK--SGDPAALPKSHTCNFRVLDPPYKSYET 851

731 LYELTALTAIBETCGFAVE 748
852 LEHKWSIAVETLGFQGE 869

RESULT 8

1-10-205-823-279

Sequence 279, Application US/10205823

Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John E.

APPLICANT: Endegé, Wilson O.

APPLICANT: Gannavarapu, Manjula

APPLICANT: Gorbacheva, Bella

APPLICANT: Hoersch, Sebastian

APPLICANT: Kamatkar, Shubhangi

APPLICANT: Wensey, Angela M.

APPLICANT: Glatt, Karen

APPLICANT: Zhao, Xumei

APPLICANT: Anderson, Dustin

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

THERAPY OF PROSTATE CANCER

FILE REFERENCE: VRI-044

CURRENT APPLICATION NUMBER: US/10/205,823

PRIOR FILING DATE: 2002-07-25

PRIOR FILING DATE: 2001-07-25

PRIOR FILING DATE: 2001-07-25

PRIOR FILING DATE: 2001-08-22

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2002-03-05

NUMBER OF SEQ ID NOS: 455

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 279

LENGTH: 911

TYPE: PRT

ORGANISM: Homo sapiens

-10-205-823-279

Query Match

Best Local Similarity 33.9%; Score 1370.5; DB 15; Length 911;

Matches 324; Conservative 119; Mismatches 234; Indels 207; Gaps 28;

24 LVKXDFRFP---DFFAKV---YVDGQGH---STDVKNLTDPKWNQHYDLYIGKSD-S 74
69 LPTSDELFLPGICDPYVVLKSLYVADENRELALVQTKTKTLNPKWNEEFYFRVNPNSHR 128
75 VTSVNVNKHKKQAGFLGCVR-----LLSNAINRLKDTGY 112
129 LFEVFDENLRTDD---FLGQVDVPLSHLPTDPTMERYTFKDFLLRPSHKS RVKG 185
113 QRLDLCKLGNNDTVRGQTVLSQSRDRIGTGGVDCSRLFDND-----L 159
186 LRLKWAYMPKNG---GQDEENSQDQDDMEHGWVD-----SNDASQHEBLPPPL 235
160 PDGWEERTASGRIOVNLNITRTQWERT-----RPA 192
236 PGWESEKVDNLGRTYVNNHNTTQWHRSLMDVSESDNNIRQINQEAHRRFRSRRI 295
193 SEYSSP-----GRFLSCFVDENPTISGTN-----GATCQSSDPR-LAE---RR 232
296 SEDLEPEPSGGVPPEWET-ISEEVNIAGDSLGLALPPFPASPSRTSPSELSELSRR 354
233 VR-----SQHRNVMGRTHL-----HTPPD----- 252
355 LQITPDNSGEQFSLIQREPFSSRLRSCSVTDAVBQGLHLPFGAKDSPVRAVKDTLSNPQ 414

QY 253 -----LPEGQERTTQGGQVYFUHTOTGVSTWHDPRVP-----RD 287
Db 415 SPQSPYNSPKQHVKVTSQFLPPGWMEMIAFNPGFFFDHNTKTTTWEDPRLKFPVHMRS 474
QY 288 LSNINCEBELGPLPGWEIRNTATGRVYFVDHNNRTQFTDRLSANLHLVLNRQNLKQD 347
Db 475 KTSLNENDLGLPLPGWEERIHLDTGRTFYIDHNSKITQWEDPRL-----QN----- 519
QY 348 QQQQVSLCPDDTECLTVPR--YKRDVLQKLIKLRQELSQQQPOQAGHCRIEVSREEIFE 405
Db 520 -----PAITGPAVPSREFPKQDYFRKKLKKPADIPNRFEMKLRHNNIFE 566
QY 406 SYRQVMKMRPKDLWK-RLMIXFRGBEGLDYGVAREWLYLLSHMLNPYGLFOYSRDDI 464
Db 567 SYXRMSVKRDPDLKARLWIEFESEKGLDYGVAWEFFLLSKEMFNPYGLFYSATDN 626
QY 465 YTLQNPDSAV-NPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLFPYKQLGKSTLDDME 523
Db 627 YTLQINPNSGLCNEDHLSYFTFGRVAGLAVFHGKLLDGGFFIRFPYKMLGKQITLNDME 686
QY 524 LYDEPDLHSLVWILENDITGVLDHTFCVHNAYGBILQHHLKPNKSIIPYNEENKKEVVR 583
Db 687 SVDSYKSLKWLILENDPT-BLDMFCIDEENFGQTYQVDDLKPNSEIMVTNENKREYID 745
QY 584 LYVNWRFLEIGIEAQFLALQKGFNEVIPOHLLKTFDEKELELIIICGLKIDVNDWKVNTRL 643
Db 746 LVIQWRFVNRVQKMNALFEGFTTELLPIDLIKIFDENELELLMCGLDGVDVNDWRQHSIY 805
QY 644 K-HCTPDPSNVKWFKAVEFPDEERRARLLQFVTGSSRVPLQGFKALQGAAPRLFTIH 701
Db 806 KNGYC-PNHPVIQWFKAVLLMDAEKRIQLLQFVTGSRVPMNGFAELYSNGPQLPTIE 864
QY 702 QIDACTNNLPKACTCNRIDIPPYSEYKLYEKLTLTAIETCGF 745
Db 865 QWGS-PEKLPRAHTCFNRLDLPYETPEDIREKLLMAVENAQGF 907

RESULT 9

US-10-185-050-126

Sequence 126, Application US/10185050

Publication No. US20030077577A1

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio

Key, Brian K.

Fowkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 233

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT NUMBER: US/10/185,050

FILING DATE: 28-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/826,516

FILING DATE: 03-Apr-1997

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 19,872

REFERENCE/DOCKET NUMBER: 1101-208-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 126:

SEQUENCE CHARACTERISTICS:

LENGTH: 725 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 126:

-10-185-050-126

Query Match

Best Local Similarity 33.6%; Score 1355.5; DB 15; Length 725;

Mismatches 280; Conservative 98; Indels 63; Gaps 13;

137 QSRDRIGTGQVDCSRLFDN---DLPDGWEERTASGRIOVLNHLTRITTTQWERPRPAS 193

112 RARSSTVTGGEPTPSVAYVHTTGLPSGWEERKDAKGRYYVNNHNRITTTWTRPIMQLA 171

194 EYSSPG-----RPLSC---FVDENTPISGTNGATCGQSSDPRLAERVRS 235

172 EDGASGATNNHLEPQIRPRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPQSPQ 231

236 QRHNTMSRTHLTPDLPPEGIEQRTQCGQVYFLHTQGVSTWHDPRVP-----RDLSN 290

232 PSPYNSPKQHKVTSQSLPFGWENRIAPNGRPFIDHNTKTITTWEDPRLKPPVHMRKTS 291

291 INCEELGPIPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQLKQDQQQ 350

292 LNPNDLGPLPFGWEERHLDGRFTYIDNSKITQWEDPRL-----QN----- 333

351 QVVSICPDDECTLVPR--YKDLVQKLRLQELSQOQQAQCHRIEVSREIPEESYR 408

334 -----PAITGPAPVYSREFKQYDFRKLKPKADIPNRFEMKLRHNNIFEESYR 383

409 QVMKMRPKDLWK-RLMIKFRGEGLDYGVAREWLVLSSHEMLNPPYGLFOYSRDDIYTL 467

384 RIMSVPKPDVLKARLWIEFSEKGLDYGVAKEWFFLLSKEMFPYGLFYATDNYTL 443

468 QINPDSAV-NPEHLSYFHFVGRIMGNMVFHGHYIDGGFTLPFYKQLGKSLTDDMBLVD 526

444 QINPNSGLCNEDHLSYFTFGRVAGLAVFHGKLDGGFFIRFFYKQWMLGKQITLNDMESVD 503

527 POLHNSLWILENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKIPVNEENKKEYVRLYV 586

504 SEYNSLKWILENDPT-ELDLMFCIDEENFGQYQVDLKPNGSEIMVTNENKREYIDLVI 562

587 NWRFLRGIEAQFLAQGFNEVIPHQLLKTDFEKELELIICGLGKIDVNDKVNTRLK-- 644

563 QWRFVNRVQKQNAFLQGFTELLPIDLKIIFDENELELLMGLGDVDVNDWRQHSIYKNG 622

645 HCTPDSNIVKFWKAVFDFEBERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQID 704

623 YC-PNHPVQWFWKAVLLMDAKRILQLQFVTGTSRVPMPNGFASLYGNGPQLFTIEQWG 681

705 ACTNNLPKATCFNRIDIPYSEKYLEKLLTALTEETCGF 745

682 S-PEKLPKATCFNRLDLPYETFDLQEKLLMAVENAQGF 721

SULT 10

-10-205-823-277

Sequence 277, Application US/10205823

Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John E.

APPLICANT: Endege, Wilson O.

APPLICANT: Gannavarapu, Manjula

APPLICANT: Gorbacheva, Bella

APPLICANT: Hoersch, Sebastian

APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 277
LENGTH: 854
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-277

Query Match 33.5%; Score 1354.5; DB 15; Length 854;

Best Local Similarity 43.7%; Pred. No. 3.2e-116;

Mismatches 280; Conservative 98; Indels 63; Gaps 13;

QY 137 QSRDRIGTGQVDCSRLFDN---DLPDGWEERTASGRIOVLNHLTRITTTQWERPRPAS 193

Db 241 RARSSTVTGGEPTPSVAYVHTTGLPSGWEERKDAKGRYYVNNHNRITTTWTRPIMQLA 300

QY 194 EYSSPG-----RPLSC---FVDENTPISGTNGATCGQSSDPRLAERVRS 235

Db 301 EDGASGATNNHLEPQIRPRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPQSPQ 360

QY 236 QRHNTMSRTHLTPDLPPEGIEQRTQCGQVYFLHTQGVSTWHDPRVP-----RDLSN 290

Db 361 PSPYNSPKQHKVTSQSLPFGWENRIAPNGRPFIDHNTKTITTWEDPRLKPPVHMRKTS 420

QY 291 INCEELGPIPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQLKQDQQQ 350

Db 421 LNPNDLGPLPFGWEERHLDGRFTYIDNSKITQWEDPRL-----QN----- 462

QY 351 QVVSICPDDECTLVPR--YKDLVQKLRLQELSQOQQAQCHRIEVSREIPEESYR 408

Db 463 -----PAITGPAPVYSREFKQYDFRKLKPKADIPNRFEMKLRHNNIFEESYR 512

QY 409 QVMKMRPKDLWK-RLMIKFRGEGLDYGVAREWLVLSSHEMLNPPYGLFOYSRDDIYTL 467

Db 513 RIMSVPKPDVLKARLWIEFSEKGLDYGVAKEWFFLLSKEMFPYGLFYATDNYTL 572

QY 468 QINPDSAV-NPEHLSYFHFVGRIMGNMVFHGHYIDGGFTLPFYKQLGKSLTDDMBLVD 526

Db 573 QINPNSGLCNEDHLSYFTFGRVAGLAVFHGKLDGGFFIRFFYKQWMLGKQITLNDMESVD 632

QY 527 POLHNSLWILENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKIPVNEENKKEYVRLYV 586

Db 633 SEYNSLKWILENDPT-ELDLMFCIDEENFGQYQVDLKPNGSEIMVTNENKREYIDLVI 691

QY 587 NWRFLRGIEAQFLAQGFNEVIPHQLLKTDFEKELELIICGLGKIDVNDKVNTRLK-- 644

Db 692 QWRFVNRVQKQNAFLQGFTELLPIDLKIIFDENELELLMGLGDVDVNDWRQHSIYKNG 751

QY 645 HCTPDSNIVKFWKAVFDFEBERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQID 704

Db 752 YC-PNHPVQWFWKAVLLMDAKRILQLQFVTGTSRVPMPNGFASLYGNGPQLFTIEQWG 810

[illegible][illegible]

```
705 ACNNNLPKANTCFNRIDIPPYESYKYELXLLTAIBETQGF 745  
:  
811 S-PEKLPRAHCTFNLDPPTETFDLREXLMAVENAQOF 850
```

SULT 11

-10-097-534-15

Sequence 15, Application US/10097534

Publication No.US20030049607A1

GENERAL INFORMATION:

APPLICANT: GREENER, TSVIKA

APPLICANT: MOSKOWITZ, HAIM

APPLICANT: REISS, YUVAL

APPLICANT: ALROV, IRIS

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL TITLE OF INVENTION: MATURATION

FILE REFERENCE: PLV-001_01

CURRENT APPLICATION NUMBER: US/10/097,534

PRIOR FILING DATE: 2002-03-12

PRIOR APPLICATION NUMBER: 60/275,224

PRIOR FILING DATE: 2001-03-12

PRIOR APPLICATION NUMBER: 60/308,958

PRIOR FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: 60/340,170

PRIOR FILING DATE: 2001-12-07

NUMBER OF SEQ ID NOS: 71

SOFTWARE: PatentIn ver. 2.1

SEQ ID NO 15

LENGTH: 927

TYPE: PRT

ORGANISM: Homo sapiens

-10-097-534-15

```
705 ACTNNLPKAHCFNRIIDIPPYESYEKLYEKLTAIETCGF 745  
      :   |||:|||||:|||||:|||||:|||||:||  
811 S-PEKLPRAHCFNRLDIPPYETFDLREKLNAVENAQGF 850
```

```
261 TTQGOVYFHTQTGVSTWHDPRV-----RDLNINCEELGPPPGWEIRNTATGRVVF 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
527 IAPNGRPFIDNTKTITWEDPRLKFPVHMSKTSLNENDLGPPLPGWEIRIHLDCRTFY 586
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
316 VDHNRRTQFTDPRLSANLHLVLRQNLKDOQQOQVSVLCPDDTECLTVPR--YKRDIV 373
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
587 IDHNSKITQWEDPRL-----QN-----PAITGPAPVYSREFK 618
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
374 QKLKILRQLSQQOQAGHCRIEVSREEIFESYRVQVMKPKDLWK-RLMIKFRGEEGL 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
619 QKYDYFRKLLKPPADIENRPFEMKLRHNNIFEESYRVRMSVKRDPVLKARLWIEFESEKGL 678
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
433 DYGGVAREWYLLSHEMLNPFYGLFOYSRDDIYTLQINPDSAV-NPEHLSYHFVGRIMG 491
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
679 DYGGVAREWFFLLSKEMFNPFYGLFEYSATDNYTLQINPNSGLCNEHDHLSYFTFIGRVAG 738
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
492 MAVFHGHYIDGGFTLPFYKOLLGKSIITLDDMELVDPDLHNSLWILENDITGVLDHTFCV 551
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
739 LAVFHGKLLDGGFFIRFPYKMKLQKQITLNDMESVDSEYNSLWILENDPT-ELDMFCI 797
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
552 EHNAYGEIIQHELKPKNGKSI PVNEENKKEYVRLVYNWRFLEAGIEAQFLAQKGFNEVPO 611
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
798 DEENFGQTYQVLDKPKNGSEIMVTNENKREYIDLVIQWRFVNRVCKQWNAFLEGFTELLPI 857
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
612 HLLKTFDEKELELIICGLGKIDVNDKVNTRLK--HCTPDSNIVKWFKAVERFDEERRA 669
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
858 DLIKFDENELELLMCGLDGVDVNDWRQHSIYKNGYC-PNHFPVQWFKAVLLMDAEKRI 916
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
670 RLLQFVTGSSRVPLQGFKALQAGAPRLETHIQIDACTNNLPKHAHTCFNRIDIPPYESYE 729
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
917 RLLQFVTGTSRVPNGFAELYGSGNGPQLFTIEQWGS-PEKLPRAHTCFNRDLDPYETFE 975
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
730 KLYEKLTAETTCGF 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
976 DUREKLLMAVENAQGF 991
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

MULT 13
10-205-823-275
Sequence 275, Application US/10205823
Publication No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endese, Wilson O.
APPLICANT: Ganavara, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Woneey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
PRIOR FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 275
LENGTH: 995
```

```
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-275

Query Match
Best Local Similarity 33.5%; Score 1353.5; DB 15; Length 995;
Matches 275; Conservative 44.6%; Pred. No. 51e-116; Indels 60; Gaps 12;

QY 159 LPDQWERRRTASGRIOVLNHTTTQWERTTRPASYSRPG-----RPLS 203
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 407 LPSGWEERKDAKGRYYVYNNRRTTTTTRPIMQLAEDGASGATNSNNHLIEPQIRPRS 466
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 C---FVDENPTISGTCATCGQSSDPELAERRVSRHRHYNMRSRTHLHTPPDLPEGEOR 260
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 467 LSSFTVLSAPLEAKDSPVRAVKOTLSNPQSPQSPYNSPKPQHKVTCQSLFPQWEHR 526
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 261 TTQGOVYFHTQTGVSTWHDPRV-----RDLNINCEELGPPPGWEIRNTATGRVVF 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 IAPNGRPFIDNTKTITWEDPRLKFPVHMSKTSLNENDLGPPLPGWEIRIHLDCRTFY 586
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 VDHNRRTQFTDPRLSANLHLVLRQNLKDOQQOQVSVLCPDDTECLTVPR--YKRDIV 373
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 587 IDHNSKITQWEDPRL-----QN-----PAITGPAPVYSREFK 618
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 374 QKLKILRQLSQQOQAGHCRIEVSREEIFESYRVQVMKPKDLWK-RLMIKFRGEEGL 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 619 QKYDYFRKLLKPPADIENRPFEMKLRHNNIFEESYRVRMSVKRDPVLKARLWIEFESEKGL 678
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 433 DYGGVAREWYLLSHEMLNPFYGLFOYSRDDIYTLQINPDSAV-NPEHLSYHFVGRIMG 491
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 679 DYGGVAREWFFLLSKEMFNPFYGLFEYSATDNYTLQINPNSGLCNEHDHLSYFTFIGRVAG 738
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 492 MAVFHGHYIDGGFTLPFYKOLLGKSIITLDDMELVDPDLHNSLWILENDITGVLDHTFCV 551
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 739 LAVFHGKLLDGGFFIRFPYKMKLQKQITLNDMESVDSEYNSLWILENDPT-ELDMFCI 797
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 552 EHNAYGEIIQHELKPKNGKSI PVNEENKKEYVRLVYNWRFLEAGIEAQFLAQKGFNEVPO 611
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 798 DEENFGQTYQVLDKPKNGSEIMVTNENKREYIDLVIQWRFVNRVCKQWNAFLEGFTELLPI 857
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 612 HLLKTFDEKELELIICGLGKIDVNDKVNTRLK--HCTPDSNIVKWFKAVERFDEERRA 669
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 858 DLIKFDENELELLMCGLDGVDVNDWRQHSIYKNGYC-PNHFPVQWFKAVLLMDAEKRI 916
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 670 RLLQFVTGSSRVPLQGFKALQAGAPRLETHIQIDACTNNLPKHAHTCFNRIDIPPYESYE 729
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 917 RLLQFVTGTSRVPNGFAELYGSGNGPQLFTIEQWGS-PEKLPRAHTCFNRDLDPYETFE 975
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 730 KLYEKLTAETTCGF 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 976 DUREKLLMAVENAQGF 991
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 14
US-10-313-955-6
Sequence 6, Application US/10313955
Publication No. US20030199036A1
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nefsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```


575 YLEPLAAEDDQKINTGPFRLVTEWKIVKRYVEQFNAFMSGFNELIPADLVNVFDEREL 634
623 ELJICGIGKIDVNDWKVNTLKHCTPDSNIVKFWKAVEFFDEERRALLQFVTGSSRVP 682
635 ELJIGGIADIDVDWKKGYDYGYESDEVIONFWKIVRSWDAEQKSLLOFTTGTSRIP 694
683 LQGFKALQGAAGPRLFTIHQIDACTNNLKFAHTCFNRIDIPPPYSEYKLYEKLJTAIEET 742
695 VNGFKDLOQSDGPRRFTIEK-SGDPAALFKSHTCFNRDLDPYKSYETLEHKMSIAVEET 753
743 CGFAVE 748
754 LGRQOE 759

arch completed: February 20, 2004, 15:37:46
b time : 36.0523 secs

56 WNETFEVNVNDNSTIAIQVFDQKFP-KKKGGQGLGVINLRVGVLDLAIGGDEMILTRDLK 114
119 KLGPNNDNTVRGQIVVSLQ-----DLPDGEERTASGRIOYLNLHITRTQ 137
115 K--SNENTVVGKIIINLSTTAOSTLQVPSSAASGARTORTSITNDPQSSQSSVSRNPA 172
138 -----SRDRIGTGGQVDCSLFDN-----DLPDGEERTASGRIOYLNLHITRTQ 184
173 SSRAGSTRDNAPASAPSEPTFFSFDQYGLPGWERTDNLGRYYIVDHTRTST 232
185 WERTP-----RPASEYSSPGRLSCFVDENTPISGTNGATCGQSSDPPLAERRVRSQRH 238
233 WIRPNLSSVAGAAAELHSSAS--SANVTGQVQSSSNA--RTEASVLT--- 279
239 RNYMSRTHLTPDLPPEGYBORTTOGQGVFLHTQGVSTWHDPRVPRDLNIN----- 292
280 ----SNATTAGSGELPGWERTPEGRPFVFDHNTRTTWDPRQQTIRSYGGPNAT 335
293 -----CEELGPLPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQQLKDQ 347
336 IQQPVSQLGPLSGWEMRLTNTARVYFVDHNTKTITWDPRLPSSL-----DQ 384
348 QQQQVSLCPDDTECLTPRYKRDVLQKILRQELSQ--QPOAGHCRIEVSREEIFE 405
385 -----NVPOYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHIIFED 425
406 SYRQVWMPKPKLWKLMIKFRGEEGLDYGVAREWLYLLSHEMLNPPYGLFOYSRDDIY 465
426 SYAEINRQSATDLKRLMIKFDGEDGLDYGLSREYFLLSHEMNFYCLFEYSVDNY 485
466 TLQINPDSAVNPEHLSPYFVGRINGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELV 525
486 TLQINPHSGINPEHLNPFKIGRVIGLAIFHRRFVDFAFFVVSFYKMLQKKVTLQDMESM 545
526 DPDLHNSLVILENDITGVLDHFCVEHNAYGBIIQHLEKPKNGKSIIPVNEENKKEYVRLY 585
546 DAERYSLVILNDITGVLDITFVEDNCFGEVVITDLKPNGRNIEVTENKREYVDLV 605
586 VNRFLRGIEAQFLAQGFNEVLPQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKH 645
606 TVWRIQRIEIQEAFNAFEGFSELIPQELINVFDERELLELIIGGISEIDMEDWKKHDTYRS 665
646 CTDPDSNIVKFWKAVPEFDEREARLLQFTGSSRVPLOGFKALOGAAGRLFTTHQIDA 705
666 YSENDQIIKFWELMDWSNEKSKRLQFTTGTSGRIIPVNGFKDLQSGDGRKFTIEKAGE 725
726 -PNKLPKATCFNRLDLPYTSKKDLHKLSTAVEETIGFGQE 767

RESULT 2
16562
ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)
Alternate names: E6-AP-like protein ubiquitin ligase
Species: Schizosaccharomyces pombe
Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
Accession: S66562, T45159
Nefsky, B.; Beach, D.
DOI: J. 15, 1301-1312, 1996
Title: Pub1 acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25
Reference number: S66562; PMID:96205868; PMID:8635463
Accession: S66562
Status: preliminary; nucleic acid sequence not shown
Molecule type: mRNA
Residues: 1-766 <NEF>
Cross-references: GB:U66716; NID:g1519443; PIDN:AAB07514.1; PID:g1519444
Nefsky, B.S.; Beach, D.
Submitted to the EMBL Data Library, August 1996
Description: Pub1 acts as an E6-AP-like protein ubiquitin ligase in the degradation of
Reference number: Z22935
Accession: T45159
Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA
A: Residues: 1-766 <NE2>
A: Cross-references: EMBL:Y07592; PIDN:CAA68867.1
C: Genetics:
A: Gene: pub1
C: Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C: Keywords: ligase
F: 205-242/Domain: WW repeat homology <WW1>
F: 288-325/Domain: WW repeat homology <WW2>
F: 345-382/Domain: WW repeat homology <WW3>

Query Match 38.7%; Score 1564; DB 1; Length 766;
Best Local Similarity 41.1%; Pred. No. 3, 7e-105; Indels 132; Gaps 18;
Matches 339; Conservative 121; Mismatches 233

QY 1 MNPGRRRNGPVKRLTLCANLVKDFRLPPFAKVVVDGSGQCHSTDTVKNITLDPK 60
DB 1 MNSAQSR-----RIRVTIVAADGLYKRDVFRFPDPFAVLTVDGE-QTHHTTAIKKTLNPF 55
QY 61 WNOHYDLYIGKSDVTISWNHKKIKKQAGAGLGCVPRLLSNAINRLKDKTCYQRL--DLC 118
DB 56 WNETFEVNVNDNSTIAIQVFDQKFP-KKKGGQGLGVINLRVGVLDLAIGGDEMILTRDLK 114
QY 119 KLGPNNDNTVRGQIVVSLQ----- 137
DB 115 K--SNENTVVGKIIINLSTTAOSTLQVPSSAASGARTORTSITNDPQSSQSSVSRNPA 172
QY 138 -----SRDRIGTGGQVDCSLFDN-----DLPDGEERTASGRIOYLNLHITRTQ 184
DB 173 SSRAGSTRDNAPASAPSEPTFFSFDQYGLPGWERTDNLGRYYIVDHTRTST 232
QY 185 WERTP-----RPASEYSSPGRLSCFVDENTPISGTNGATCGQSSDPPLAERRVRSQRH 238
DB 233 WIRPNLSSVAGAAAELHSSAS--SANVTGQVQSSSNA--RTEASVLT--- 279
QY 239 RNYMSRTHLTPDLPPEGYBORTTOGQGVFLHTQGVSTWHDPRVPRDLNIN----- 292
DB 280 ----SNATTAGSGELPGWERTPEGRPFVFDHNTRTTWDPRQQTIRSYGGPNAT 335
QY 293 -----CEELGPLPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQQLKDQ 347
DB 336 IQQPVSQLGPLSGWEMRLTNTARVYFVDHNTKTITWDPRLPSSL-----DQ 384
QY 348 QQQQVSLCPDDTECLTPRYKRDVLQKILRQELSQ--QPOAGHCRIEVSREEIFE 405
DB 385 -----NVPOYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHIIFED 425
QY 406 SYRQVWMPKPKLWKLMIKFRGEEGLDYGVAREWLYLLSHEMLNPPYGLFOYSRDDIY 465
DB 426 SYAEINRQSATDLKRLMIKFDGEDGLDYGLSREYFLLSHEMNFYCLFEYSVDNY 485
QY 466 TLQINPDSAVNPEHLSPYFVGRINGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELV 525
DB 486 TLQINPHSGINPEHLNPFKIGRVIGLAIFHRRFVDFAFFVVSFYKMLQKKVTLQDMESM 545
QY 526 DPDLHNSLVILENDITGVLDHFCVEHNAYGBIIQHLEKPKNGKSIIPVNEENKKEYVRLY 585
DB 546 DAERYSLVILNDITGVLDITFVEDNCFGEVVITDLKPNGRNIEVTENKREYVDLV 605
QY 586 VNRFLRGIEAQFLAQGFNEVLPQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKH 645
DB 606 TVWRIQRIEIQEAFNAFEGFSELIPQELINVFDERELLELIIGGISEIDMEDWKKHDTYRS 664
QY 646 CTDPDSNIVKFWKAVPEFDEREARLLQFTGSSRVPLOGFKALOGAAGRLFTTHQIDA 705
DB 666 YSENDQIIKFWELMDWSNEKSKRLQFTTGTSGRIIPVNGFKDLQSGDGRKFTIEKAGE 724
QY 706 CTNNLPKATCFNRLDIPYTSKKDLHKLSTAVEETIGFGQE 748
DB 725 -PNKLPKATCFNRLDLPYTSKKDLHKLSTAVEETIGFGQE 766

RESULT 3

49744
robable ubiquitin-protein ligase [imported] - Neurospora crassa
;Alternate names: protein B24B19.160
;Species: Neurospora crassa
;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
;Accession: T49744
;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
;Reference number: Z25022
;Status: preliminary
;Molecule type: DNA
;Residues: 1-815 <SCH>
;Cross-references: EMBL:AL556192; GSPDB:GN00116; NCSP:B24B19.160
;Experimental source: BAC clone B24B19; strain OR74A
;Genetics:
;Gene: NCSP:B24B19.160
;Map position: 6
;Introns: 11/1; 24/1; 59/2; 110/1; 783/2
;Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
;239-276/Domain: WW repeat homology <WW1>
;334-371/Domain: WW repeat homology <WW2>
;393-430/Domain: WW repeat homology <WW3>
Query Match 37.7%; Score 1521; DB 2; Length 815;
Best Local Similarity 38.8%; Pred. No. 5.3e-102; Mismatches 229; Indels 170; Gaps 20;
Matches 332; Conservative 124; Mismatches 229; Indels 170; Gaps 20;
16 LTVCACNLVKDFFRDPDPFAKVVVDSGQCHSTDTVKNLDPKKNQHYDLY-----I 69
9 LEVIAADGLYKRDVERPDPFAVATNGE-QTKTQVSKRTLPYWNHFDPLQORSKV 67
70 GKSDSVTSVNNKKHKQAGLGVCL-LSNAINRLKDTGYQL--DLCKLGNND 126
58 NEDSILAVQFDQKF-KKKDQGLGVINRIGVIELAPDAEDMLTRDLK--STDNL 124
127 TVRQGVVSLQ-----SRDRI-----GTG 145
125 VVHGKLIINLTNLATMSRLGPPSPSSRSLTTPQSSVINSDRANERSSAMSGNGTA 184
146 GOVVDCSR-----LFDND---LPDND 162
185 NMTLARSPLASVSSSTAPTPTNGTAPTNPSTLTPAQARHHSSTLSPFSDSMGRLPAG 244
163 WEERTASGRIQYLNHTRTQWERPRPASEYSSPGRPLSCFVDENTPI----- 212
245 WEREDHLGRYYVDHNSRTSNRPT-----GTGAENRTAENTQVERQQRNRTL 297
213 ----SGNGATCGOSSDPRLAERRVRSQRHNYSRTHLTP--PDLPEGYQRITQQGQ 266
298 PEDRTGAN-----SPTLQQQQAATAATAMHTGATTPTGELPAGWEQRTPPGR 349
267 VYFLHTQTGVSTHDPVRPDL-----SNIN-----CEELGLPPEWEIRNTATGRV 313
350 PYFDVHNRITTTWDPRRQYIRMYGQNNNTGTIOQPVSQLGLPLSGWENRLNTARV 409
314 YFVDHNRITTTQFDPLRLSANLHLVLRNQLKQDQOQVVSCLPDDTECLTVPRYKRLV 373
410 YFVDHNRITTTWDPRLPSSL-----DQ-----NVQYKRDGR 442
374 QKLKILACELSSQOQAGHCRIVSREEIFESYRVQVMQKRPDKLWRLMKFRGEGLD 433
443 RLLYFRSQ-PAMRIMGQCHIKVRSHIFEDSFABEISQASATDLKXLMIKFDGEDGLD 501
434 YGGVAREWLYLLSHEMLNPPYGLFQYSRDDIVTLQINPDSAVNPEHLSVFHFVGRMGMA 493
502 YGGLSREFFLLSHEMFNPPYCLFEYSADHNTYTLQINPHSGINPEHLYFFYIGRVGLA 561
494 VFCHYIDCGFTLFFPKQLLGKSIITLDDMELVDPDLHNSLVMLINDITGVLDHTFCVEH 553
562 IFHRFDLDAFFIGALYKMWLGKAVSLADMEGVADPFRSLQWMLDNDITDVLDAFTFSD 621
554 NAYGEIITQELKPKNGKSIIPWNEENKKEYVRLVYVNRFLRGIEAQFLAQKGFNFIQHL 613

Db 522 ERFGVITEEDLPLNGRNIAVTNENKKEYVELMVKWAKRIEQQFRAFQDGHFELIPOOL 681
QY 614 LKTDEKELELIICGLGKIDVNDKVNTRLKHCTPDNSIVKWFKAVFEPDEERRARLIQ 673
Db 682 INVDERELBELLIGIAEIDVDKKKHTDYRGYTESDEVIOFFWQTVRSWDGEGQKSRLLQ 741
QY 674 FVTGSSRVPLQGFKAQAGPRLFTIHOIDACTNVLPKATCFNRIDIPPYESVEKLYE 733
Db 742 FTGTISRIPVNGFDLQSGDGPFRFTIERAGEIT-NLPKAHTCFNRLDLPFYKSLMLQQ 800
QY 734 KLLTAIEETCGFAVE 748
Db 801 KLLTAIEETMGFQGE 815
RESULT 4
S43217
ubiquitin-protein ligase (EC 6.3.2.19) RSPs - Yeast (Saccharomyces cerevisiae)
N/Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein
C/Species: Saccharomyces cerevisiae
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C/Accession: S43217; S50628; S70050
C/Mulligan, J.R.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor
submitted to the EMBL Data Library, February 1993
A/Reference number: S30812
A/Accession: S43217
A/Molecule type: DNA
A/Residues: 1-809 <MUL>
A/Cross-references: GB:U18916; EMBL:J111119; NID:G1384128; PIDN:AAC03223.1; PID:G603364
R/Dietrich, F.S.
A/Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and 1ambc
A/Reference number: S50628
A/Accession: S50628
A/Molecule type: DNA
A/Residues: 1-809 <DIE>
A/Cross-references: EMBL:U18916; NID:G1384128; PIDN:AAC03223.1; PID:G603364; GSPDB:GNC
R/Hein, C.; Springael, J.Y.; Volland, C.; Hagenauer-Teapls, R.; Andre, B.
Mol. Microbiol. 18, 77-87, 1995
A/Title: NP11, an essential yeast gene involved in induced degradation of Gap1 and Fur
A/Reference number: S70050; MUID:96154942; PMID:8596462
A/Accession: S70050
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-101 <HEI>
C/Genetics:
A/Gene: SGD:RSP5; PUB1; NP11; MIPS:YER125w
A/Cross-references: MIPS:YER125w; SGD:S0000927
A/Map position: 5R
C/Function:
A/Description: Involved in endocytosis of GAP1 protein and FUR4 protein; binds and ubi
C/Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C/Keywords: ligase
F/229-266/Domain: WW repeat homology <WW1>
F/331-368/Domain: WW repeat homology <WW2>
F/387-424/Domain: WW repeat homology <WW3>
Query Match 36.8%; Score 1484; DB 1; Length 809;
Best Local Similarity 37.1%; Pred. No. 2.5e-99;
Matches 324; Conservative 125; Mismatches 224; Indels 200; Gaps 20;
QY 11 PVKRLTLVLCANLVKDPFRLPDPAKVVVDSGQCHSTDTVKNLDPKKNQHYDL-YI 69
Db 2 PSISVKLVAESLYKRDVFRSPDPFAVLTIDGY-QYKTSAAKTLNLPYWNHFFKFDI 60
QY 70 GKSDSVTSVNNKKHKQAGLGVCL-LSNAINRLKDTGYQL-TCYQR-----LDLC 118
Db 61 NENSILITQVFDQKF-KKKDQGLGVINRIGVIELAPDAEDMLTRDLK 119
QY 119 KLGPNDNDTVRQIVV-----SLSQRDRIGTGGQVDCSR--- 153
Db 120 K--SNDGMVSGRLIVLKLSPSSPHSQAPSGHTASSTNTSTTTNGHST--SSTNHS 177

154 -----LEND-----LPDCHEE 165
178 TSPSRGTAGAVSTLQSGTAAATNTATNTSHRSTNSTSSATRYQSFEDQYGRUPPGWER 237
166 RTTAGRIQYLNHITRTTOWERTRPASE-----YSSGRPLSCFVDENT 210
238 RTDNFGRYYVDHNTRTTWTWRPLDQTEAERGQNLNANTELERQHRGRTLPGGSSDNS 297
211 -----PISG-----TNGATCQSSDPRLAERVSQRHRYMSRTHLHT 249
298 SVTVQVGGGNIIPVNGAAAAFAATGGTSGLG----- 331
250 PPDLPEGYEORTTQCGQVFLHTQTGVSTWHDPVRPRDLNIN-----CBELQPL 299
332 --ELPSGWEGRFTPEGRAYFVDHNTRTTWTWDPRRQYRTYPTNTTIQQOPVSQLGLF 389
300 PGWEIRNTATGVIYVDHNTRTTQTPRLSANLHLVLRQNLKQOQQQVVSCLPDD 359
390 PSGWEMELTNTARVYVDHNTKTITWDDPRLPSSL-----DQ----- 426
360 TECLTVRYKRDVLQKILRQELSQOQFQ-----AGHCRIEVSREIEPESYRQVMKWRP 415
427 ----NVFYKRDPRKVIYFR-----SQPALRLTPQCHIKYRKNIIFEDAYQEMRQTP 477
416 KDLKRLMIKFRBEGLDYGVAREWLYLLSHMLNPPYGLFQYGRDDIYTLQINPDSAV 475
478 EDLKKRLMKFRBEGLDYGVAREWLYLLSHMLNPPYGLFQYGRDDIYTLQINPDSAV 475
476 NPEHLSYFHVGRIMGMAVPHGHYIDGGFTLPFYKOLGKSIITLDDMELVDPDLHNSLAW 535
538 NPEHLSYFHVGRIMGMAVPHGHYIDGGFTLPFYKOLGKSIITLDDMELVDPDLHNSLAW 535
536 ILENDINGVLDTFSAEDDERFGEVVTDLKPGRNIEVTDGNKKEYLYTQWRIQVDRVQ 657
598 MLENSIDGVLDTFSAEDDERFGEVVTDLKPGRNIEVTDGNKKEYLYTQWRIQVDRVQ 657
596 AQFLALQKGNVPIPHLLAKTDEKELELIICGLKIDVNDKVNTRLKHTCPDPSNVKW 655
658 EQKAFNDGNEILPDLVTVFDERELELLIGIAIEDIEDKKHTDYGQESDEVIQW 717
656 FWKAEFFDEERRARLQFVTGSSRVPLOQFKALOGAAGPRFTTHQIDACTNNLPKHAHT 715
718 FWKAEFFDEERRARLQFVTGSSRVPLOQFKALOGAAGPRFTTHQIDACTNNLPKHAHT 715
716 CFNRIDIPPPYESYKELYKLLTAIBETCGFAVE 748
777 CFNRIDIPPPYESYKELYKLLTAIBETCGFAVE 748

ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)
Species: Schizosaccharomyces pombe
Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
Accession: T39585
Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, August 1997
Reference number: Z21865
Accession: T39585
Status: preliminary; translated from G3/EMBL/DBJ
Molecule type: DNA
Residues: 1-786 <VOL>
Cross-references: EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GN00067; SPDB:SPBC16E9.11c
Experimental source: strain 972h-; cosmid c16E9
Genetics:
Gene: SPDB:SPBC16E9.11c
Map position: 2
Introns: 60/2
Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
236-273/Domain: WW repeat homology <WWR1>
306-343/Domain: WW repeat homology <WWR2>
364-401/Domain: WW repeat homology <WWR3>

Query Match 35.1%; Score 1418; DB 2; Length 786;
Best Local Similarity 36.7%; Pred. No. 1.5e-94;
Matches 316; Conservative 127; Mismatches 217; Indels 202; Gaps 20;
QY 10 GPVKRLVTLCAKLNKKKDFRLPPPAKVVVDGSGGCHSTDTVKTLDPKMNOHYDLYI 69
DB 4 GAKRVRFIVAADGLSKEDLFRQDPDFALLTVGGE-QTHITTKVKSVPNWEGEFTV 62
QY 70 GKSDSVTISVNNHKKHKKQAGFLGCVRLLSNAINRLKDTGYORLDL-----CKLGP 122
DB 63 KPSSVISRLFDQKXF-KKQDQGLGLVWF-----RMREVGSFRSNREVSLTRPLKSS 115
QY 123 NDNDTVRGQIVV----- 134
DB 116 TTNLVSLGNLVLKAPSKIRAPAGNHSSTANRTSTPTTTTARTTTRPTATTNIN 175
QY 135 -----SLOSRDRIGTG-----QVVDCSRLLFND----- 158
DB 176 QSTNSTENGTSAAATNGTGTGAGTGAHRSSPVTNRQTNTSALSNNAHIMSSPEDQY 235
QY 159 --LPDQWEERTASGRICVLAHITRTTOWERPTPASEYSSPGRPLSCFVDENTPISGTN 216
DB 236 GRLLPGWERRADSLGRTYVDHNTRTTW-----TPAS-----STNPVHT- 277
QY 217 GATCQSSDPRLAERVSQRHRYMSRTHLHTPP-----DLPEGYEORTTQ 263
DB 278 -----SSD-----SQRLNHNQNR-----HLPDDSNPSLMQSDGNDLPFGWEMRYTD 318
QY 264 QCGVYFLHTQTGVSTWHDP-----VPRDLNINCEELGRLPPGWEIR 306
DB 319 TGRPFVDHNTRTTWTWDPDRNPLVRPNGSGSTVGLMQPOSLSH-----LGPDPFGWEMR 373
QY 307 NTATGRVYVDHNTRTTQTPRLSANLHLVLRQNLKQOQQQVVSCLPDDTECLTVP 366
DB 374 LNSARVYVDHNTKTITWDDPRLPSAL-----DQD-----VP 406
QY 367 RYKRLVOKLILROELSQOQPAQHCRIEVSREIEPESYRQVMKWRPKDLKRLMIK 426
DB 407 QYKCDPRKLIYFRSQ-PGMRPLPGQCNVYKVRDHIPEDSYAEIMRYSAHDLKRLMIRF 465
QY 427 REEGLDYGGVAREWLYLLSHMLNPPYGLFQYGRDDIYTLQINPDSAVNPEHLSYFHV 486
DB 466 DEEDGLDYGGLSRERFFLLSHKMFDPYCLFYSADVNTYTLQINPHSSINPEHNTFRFI 525
QY 487 GRIMGMAVPHGHYIDGGFTLPFYKOLGKSIITLDDMELVDPDLHNSLWILENDITGVLD 546
DB 526 GRVIGLAIFFHRFLDAPFVSLYKLLRKKVSLADMESIDAIFYRSLKWLENDITGILD 585
QY 547 HTFCVHNAYGEIIQHELKPNKSIIPVNEENKKEYVLYVNRFLRGIEAQFLALQKGN 606
DB 586 LTFVEEDHFGVEVRTVELITNGENIEVTEENKKYVDLVTEWRVSKRVEQQFNAYSGFV 645
QY 607 EVIPOHLAKTDEKELELIICGLKIDVNDKVNTRLKHTCPDPSNIVKFWKAVEFFDEE 666
DB 646 ELVSPDLVNVFDERELELLIGISDVVDKWSHTEYRTVIATDPVIKWFWEIIAGWNE 705
QY 667 REARLLQFVTGSSRVPLOQFKALOGAAGPRFTTHQIDACTNNLPKHAHTCFNRIDIPPE 726
DB 706 DRKLLQFATGTSRIPVNGFRDLQSGDGRKRTIEKA-GTPDQLPVAHTCFNRLDLPDYP 764
QY 727 SYEKLKELYKLLTAIBETCGFAVE 748
DB 765 SKDTLHEKLSLAVENTVGFQNE 786

RESULT 6
T46412
ubiquitin-protein ligase (EC 6.3.2.19) NEBDD4 - human (fragment)
N:Alternate names: hypothetical protein DKFP434P2422.1
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
C:Accession: T46412

Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
Reference number: Z23034
Accession: T46412
Status: Preliminary
Molecule type: mRNA
Residues: 1-820 <AAA>
Cross-references: EMBL:AL137469
Experimental source: adult testis; clone DKFp434P2422
Genetics:
Gene: GDB:NEDD4
Note: DKFp434P2422.1
Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
Keywords: ligase
142-179/Domain: WW repeat homology <WR1>
342-379/Domain: WW repeat homology <WR2>
393-430/Domain: WW repeat homology <WR3>
489-814/Domain: ubiquitin-protein ligase homology <UB1>

Query Match 33.8%; Score 1364.5; DB 2; Length 820;
Best Local Similarity 37.1%; Pred. No. 1.2e-90;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;
50 TDTVNTLDPKMNCHYDLYGKSD-SVTISVNNHKKHKQAGFLGCVR----- 98
12 TKTIKTLNPKWNEEFVRVNPSPNHRLLFEVDENLRTDD---FLGQVDVPLSHLPTED 68
99 -----LLSNAINRLKDTGYORLDLCKLGPNDNDTVRGQVWLSQGRDRIGTGG 146
69 PTMERPTFKDPLLRPSRSHKSRVGLRWKAYMPKNG-----GQDEENSQDQDMEHW 123
147 QVDCSLFND-----LPDGHERTASGRIQYLNHTRTOWERPT----- 189
124 EVD-----SNDASQHQBELPPPLPGWEEKYDNLGRYYVNNHNRRTQWRPSLMDV 178
190 -----RPAGEYSSP-----GRPLSCFVDENTPIGTN----- 216
179 SSESNNIRIQNQAHRPRSRHSIEDLEPESGQVPEWET-ISEVNIAGDSL 237
217 -----GATCQSSDPR-LAE---RRVR-----SQRHNTYMSRTHL----- 247
238 LALPPPASPGRSRTSPQELSELSRLQITPDSNGEQFSLIQRPSSRLRSCSVTDAV 297
248 ---HTPPD-----LPGYQRTQOQ 266
298 EQGLHFGKADSIVRAVKTLNFPQSPQSPNSPKQHKVTQSLPFGWEMRIAPGR 357
267 VYFHTGTGVSTWHDPRVP-----RDLNINCEBELGFLPPGWEIRNTATGRVYFVDHNR 321
358 PFFIDHNTKITTWEDPLKFPVHMRSKTSLNPNLGLPLPGWEIRHLDGRTFYIDHNSK 417
322 TQFTDPLRLSANLHLVLRNQLKQDQOQVSVLCPDDTECLTVPR--YKRLVQKLKIL 379
418 ITQWEDPL-----QN-----PAITGPAVPSREBFKQIDYF 449
380 RQELSQOQOQAGHCHREIVSREEIFEESYRQVMKMRKDLWK-RLMKRGEGLDYGVA 438
450 RKLLKPADIPNRPEMKLHRNIFEESYRIMSKPDLVKARLATEFESEKGLDYGVA 509
439 REMLYLLSHEMLNPNYGLFOYSRDDYITQINPDASV-NPEHLSYFHFVGRIMGAVFHG 497
510 REWFLLSKEMFNFPYGLFYSATDNTYTLQINPNSGLCNEDHLSYFTFTGRVAGLAVFHG 569
498 HYIDGGTLPFYQLKGLSKITLDMELVDPDLHNSLWILENDITGLVDHTPCVEHNAYG 557
570 KLLGGFIRFPYKWLKQKITLNDMESVSEYNSLKNILENDPT-ELDLMECIDENFG 628
558 BIIQHELKPKNGKIPVNEENKKYVLYVNNRFLRGIEAQFTLALQGFNEVIFQHLKATP 617
629 QTYQVDLKPNQSEIMVTNENKREYIDIVQWRVNRVQKMAFLQEGFTLLPIDLIKIF 688
618 DEKELELIICGLKIDVMDKVNTRLK--HCTPDNSIVVMKAVKAYFFDEBRARLLQFV 675

Db 689 DENELELLMCGLDGVNDVNRQHSIYKNGYC-PNHVPIQWFKAVLLMDAEKRIQLQV 747
QY 676 TGSSRVLPQGFKALQGAAGPRLEFTIHIQIDACTNNLPKHAHTCFNRIDIPPEYSEYKLEK 735
Db 748 TGTSRVPMNGFAELYSGNGPQLFTIEQMS-PEKLPRANTCFNRLDLPYETFEDELREKL 806
QY 736 LTAIEETCGF 745
Db 807 LMAVENAQQF 816
RESULT 7
ubiquitin ligase Nedd4 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000
C:Accession: S70642
R:Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.
EMBO J. 15, 2371-2380, 1996
A:Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+)
A:Reference number: S70642; MUID:96221297; PMID:866844
A:Accession: S70642
A:Status: Preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-887 <STA>
A:Cross-references: EMBL:U50842; NID:gl293646; PIDN:AAB48949.1; PID:gl293647
C:Genetics:
A:Gene: Nedd4
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiq
F:54-167/Domain: protein kinase C2 region homology <KC2>
F:246-283/Domain: WW repeat homology <WW1>
F:402-439/Domain: WW repeat homology <WW2>
F:459-496/Domain: WW repeat homology <WW3>
F:555-881/Domain: ubiquitin-protein ligase homology <UB1>

Query Match 33.6%; Score 1355.5; DB 2; Length 887;
Best Local Similarity 37.1%; Pred. No. 6e-90;
Matches 314; Conservative 131; Mismatches 248; Indels 153; Gaps 25;
QY 14 LRLTVLCAKLVKQDFRLPDPFAKV---VDGSGQCHSTDTVNTLDPKMNQHYDLY- 68
Db 77 VRKVIAGTGLAKDILGASDPYRVLYDPMGSLTVQTKTKSLNPKWNEI-LFR 135
QY 69 -ICKSDSVTISVNNHKKHKQAGFLGCVR-----LISNAIN 105
Db 136 VLPQQRHILFEVDENLRTDD---FLGQVDVLYLPTEPRMERPTFKDFVLHPRSH 192
QY 106 RLKDTGYQRLDCKLGPNDNDTVRGQIVVLSQSRDRIGTGGQVD---CSRL-----FDN 157
Db 193 KSRVKGYLKMTYLPKNGSD---ENADQAELEBPGWVLDQPDAAATHLQHPPEPS 246
QY 158 DLPDGWEERTASGRIQYLNHTRTOWERPT-----RPAS 193
Db 247 PLPPGWEERQDVLGRYYVNNHESRTQKRPSPEDDLTDDENGDIQLQAHGAFTRQIS 306
QY 194 E-----YSSGRPLSCFVDENTPIGTNATCGQSSDP-----RLAE- 230
Db 307 EDVGDGDNHSPENWEIVREDENTYSG-QAVQPPSGHPDVQVRLAEELDTLRTMYGNP 365
QY 231 ---RRVRSQRH---RNYMSRT-----HLHTPDLPEGYEORTTQOQVYFLHTQT 274
Db 366 ATSQPTSSNHSRSGSSQTCTFEQPTLPVLLPTSSGLPPGWEKQDDGRSYVDHNS 425
QY 275 GVSTWHD-----RVP---RDLNINCEBELGFLPPGWEIRNTATGRVYFVDHNRRT 323
Db 426 KTTWSKPTMODDPRSKIPAHLRGKTPVDSNDLGLPLPGWEERTHTDGRVFFINHTKKT 485
QY 324 QFTDPLRLSANLHLVLRNQLKQDQOQVSVLCPDDTECLTVPRYKDLVQKLKILQEL 383
Db 486 QWEDPM-----QNVALTGAEP-----YSRDYKRYEFFFFRKL 519
QY 384 SOQPOAGHCHREIVSREEIFEESYRQVMKMRKDLWK-RLMKRGEGLDYGVAREW 442

520 KQOTDIPNKFEMKLRANILEDSYRRINGVKRADFLKARLWIEFDGKGLDYGVAREWF 579
443 YLLSHEMLNPPYGLFOYS-RDDIYTLQINPDSAV-NPEHLSYFHFVGRMGMAVEGHYI 500
590 FLISKEMFNPPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFKFGRVAGNAVHGLK 639
501 DGGTLPYKQKLOKSLTDDMELVDPDLHNSLWILENDITGVLDHTFCVBNHAYGEII 560
640 DGFIRFPYKQKLOKSLTLDHMSVDSYSSLRWILENDPT-ELDLRFIIDEELFGQTH 698
561 QHCLKPNGKSTPVNEENKKEYVRLVYVNRFLRGIEAQFLALOKGENEVIPOHLLKTFDEK 620
699 QHCLKTGSSEVVTNKKKXYIYVQWRFVNRIOQMAAFKGFELIPQDLIIKIFDEN 758
621 ELELIICGLKIDVNDKVNTRLKH-CTPDSNIVKWFKAVFDFDEERARLLOFVTGSS 679
759 ELELLMCGLDGVDNDWREHTKYKNGYSNLHQVHFWKAVLMDSEKRIQLLOFVTGTS 818
680 RVPLOQFKALOGAAGPRLFTIHOIDACTNNLPKATCFNRIDIPPEYSEKLYEKLTAI 739
819 RVPNGFAELYGSGNGPSQFVTEQW-GTPDKLPRAHTCFNELDLPPEYFDELWDKLOMAI 877
740 EETCGF 745
878 ENTQGF 883

SULT 8
3196
DB-4 ORF - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
Accession: I83196
Kumar, S.; Tomooka, Y.; Noda, M.
Chem. Biophys. Res. Commun. 185, 1155-1161, 1992
Title: Identification of a set of genes with developmentally down-regulated expression
Reference number: I60167; MUID:92328780; PMID:1378265
Accession: I83196
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
Residues: 1-708 <RES>
Cross-references: GB:D10714; NID:9220508; PID:9220509
Genetics:
Gene: NEDD-4
Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
40-77/Domain: WW repeat homology <WW1>
196-233/Domain: WW repeat homology <WW2>
251-288/Domain: WW repeat homology <WW3>
347-682/Domain: ubiquitin-protein ligase homology <UB1>

Query Match 29.1%; Score 1177; DB 2; Length 708;
Best Local Similarity 40.4%; Pred. No. 3.7e-77;
Matches 260; Conservative 87; Mismatches 178; Indels 118; Gaps 17;
159 LPDGWEERTASGRIOYLNHITTTQWERPT-----RPAE 194
42 LPEGWEERQVLTGRYVNHESRRQWKRPSPDDLTDNDMDQLOAQAFTTRQISE 101
195 -----YSSGRPLSCFVDENTPISG-----TNGATCGQ--S 223
102 DVDGPDNRESPEWIEVDENTVSGQVSPSGHIDVOTHAENFTRLAVCGNPAT 161
224 SDPLAERVRSORHNYMSRTH-----LHTPPDLPEGVEQRTQQGVYFL 270
162 SQP-----VTSNHSRGSGSLQTCIFBEQPTLPVLLPTSSGLPGWEEKQDRGRSYV 215
271 HTQGVSTVWHDP-----RVPRDL-SNINCEELGPPGWEIENATATGVVVDNHR 321
216 DHNSKTTTWSKTMQDDPRSKIPAHLRKGTISNDLGLPLPGWEERTTDDGRVFINNIK 275
322 TTQFTDPRLSANHLVLRNQNLKDCQQQVSLCPDDTECLTPRYKRLVOKLILRQ 381
276 KTQWEDPRL-----QNVAITGP-----AVP-YSRDYKRYBEFFR 309

382 ELSQQQPQACHRIEIVSRREEIPESYRQVMKMRPKDLWK-RLMIKFRGEEGLDYGVARE 440
310 KKKQOTDIPNKFEMKLRANILEDSYRRINGVKRADFLKARLWIEFDGKGLDYGVARE 369
441 WLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRMGMAVEGHY 499
370 WFLISKEMFNPPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFKFGRVAGNAVHGLK 429
500 IDGGTLPYKQKLOKSLTDDMELVDPDLHNSLWILENDITGVLDHTFCVBNHAYGEI 559
430 LDGFFIRFPYKQKLOKSLTLDHMSVDSYSSLRWILENDPT-ELDLRFIIDEELFGQT 488
560 IQHCLKPNGKSTPVNEENKKEYVRLVYVNRFLRGIEAQFLALOKGENEVIPOHLLKTFDE 619
489 HQHCLKTGSSEVVTNKKKXYIYVQWRFVNRIOQMAAFKGFELIPQDLIIKIFDE 548
620 KELELIICGLKIDVNDKVNTRLKH-CTPDSNIVKWFKAVFDFDEERARLLOFVTGTS 678
549 NELELLMCGLDGVDNDWREHTKYKNGYSNMHQQVHFWKAVVMDSEKRIQLLOFVTGT 608
679 SRVPLQGFKALOGAAGPRLFTIHO-----IDACTNNLPK--AHTC 716
609 SRVPMNGFAELYGSGNGPSQFVTEQWGTLLSCQETPASIAWTC 651

RESULT 9
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T37900
R;Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21752
A:Accession: T37900
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-671 <R1E>
A:Cross-references: EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:SPAC1805.15C
A:Experimental source: strain 972h-; cosmid c1805
C:Genetics:
A:Gene: SPDB:SPAC1805.15C
A:Map position: 1
A:Intons: 60/2, 105/1, 639/2
C:Superfamily: WW repeat homology
F;242-279/Domain: WW repeat homology <WWR>

Query Match 25.9%; Score 1046.5; DB 2; Length 671;
Best Local Similarity 33.9%; Pred. No. 1e-67;
Matches 257; Conservative 124; Mismatches 263; Indels 113; Gaps 17;
13 KRLITVLCANLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVKNLTPKWNQHYDLYIGKS 72
7 EVQLTILHVEGLWKNGLSLKPYLLISVD-DDQFIKTNVASGTLSLSGFTQKLVSPQ 65
73 DSVITSVNHHKHKQAGFLGCVLLSNA---INELKDTGYQRLLDLCKLGNDDNVTYR 129
66 SIILLQPLDEKQ-KNETSDGFLGAAVNSFLPFNNPKDDYKTRITL----RSPGGSYR 120
130 GQIVVVSQSRDRIGTGGQVVD---CSRFLDNDLPDGBERRTASGRIOYLNHITRTTQW 185
121 GS-VVCLPKRSKPLPEELPADKSGICTDIIDDSAGCAWETRIDFGHVYLL-----170
186 ERPTPASEYSSRPLSCFVDENTPISGTINGATCQSSDPLAERVRVSQRHNTMST 245
171 -----KSP-----QLSVISAISHEKLENLTPKQLE--VFSQFLFNNQSKS 209
246 HLHTPPD-----LPEGYEORTTQQGVYFLHTQTGVSTVWHDPVPRVPRDLNINCEELGFL 299
210 SLKINLEYKVKHLEHYPALSVRQV-----AVEKGL 244
300 PPGWEIRNTATGRVYVDHNNRTTQFTDPR-----LSANHLVLRNQNLKDCQQQ 351

Submitted to the EMBL Data Library, June 1998
Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.
Reference number: Z14334
Accession: T01491
Status: translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-1126 <YIS>
Cross-references: EMBL:AC003671; NID:G2833627; PID:G3176690; GSPDB:GN00059; ATSP:F1707
Experimental source: cultivar Columbia
Genetics:
Gene: ATSP:F1707.15
Map position: 1
Introns: 118/3; 224/3; 292/1; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2
Superfamily: ubiquitin-protein ligase homology
756-1120/Domain: ubiquitin-protein ligase homology <UBI>
Query Match 19.9%; Score 805.5; DB 2; Length 1126;
Best Local Similarity 41.3%; Pred. No. 6.5e-50;
Matches 177; Conservative 63; Mismatches 142; Indels 47; Gaps 7;
363 LTVPRYKRDVOKILRQELSSQOQPQ--AGHCRIEVSREEIFEESYRQVMKRPDLWK 420
700 LKAPRL-IDFDNKKAYFRSRIHQHCHISGPLRISVRRAYVLEDSYNQLMRSPQDLKG 758
421 RLMIKFRGEGLDYGVAREWMLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
759 RLNVQFQGEIGIDAGLTREWYQLLSRVIPDKGALLFTTVGNDATFQPNPNSVYQTEHL 817
481 SYFHFVGRIMGVAFHGHYIDGGFTLPFYKQLLGKSIITDDMELVDPDLHNSLWILEND 540
818 SYFKFVGRMVAKALFDGQLLDVYFTRSFYKHILGVKTVYHDIKAVDPDYKULWILEND 877
541 ITGVLDHTFCVE-----HNAY--GEIIQHELKPKNGKSIIPVNEENKEYVRLVYNWRFURG 593
878 VSDLDLTFSDMADEKHLIYKTEVRLMCFPCFFWCPIPKXCHICIELIILSLMKKVTDY 937
594 IEAQFLALQKGFNEVIPHLLKTFDEKELELIICGLKID----- 633
938 IRPQINAFLEGLNELPRELVSFNDEKELELIISGLPEIDCKLAFYFIFLHFAYSFKV 997
634 -----VNDKWNTRLKHCTPDNSIVKFWKAVEFDEERARLQFVTGSSR 680
998 IITILSGVFCEFLVDDLKANTETSYTSGVSPVIRFWFVWVKAFSKEDMARELFQFVTGTSK 1057
681 VPLQGFKALQGAAGPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPYESYKELYKLLTAIE 740
1058 VPLEGFKALQGISGQRLQIHKAYGSPRLPSAHCTFNQDLDFEYQSKQVQERLLAIH 1117
741 ET---CGFA 746
1118 EANEFGGFA 1126

SULT 13
6599
Olein F14J16.10 [imported] - Arabidopsis thaliana
Species: Arabidopsis thaliana (mouse-ear cress)
Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
Accession: H96599
Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
hin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
sen, N.F.; Hughes, B.; Huizar, L.
ture 408, 816-820, 2000
Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Maiti, R.; Marziali,
zzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
I. M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
Reference number: A86141; MUID:21016719; PMID:11130712
Accession: H96599
Status: preliminary
Molecule type: DNA

A;Residues: 1-4056 <STO>
A;Cross-references: GB:AE005173; NID:G9778329; PIDN:AAF79338.1; GSPDB:GN00141
C;Genetics:
A;Gene: F14J16.10
A;Map position: 1
Query Match 19.9%; Score 803; DB 2; Length 4056;
Best Local Similarity 40.0%; Pred. No. 6.9e-49;
Matches 177; Conservative 62; Mismatches 143; Indels 60; Gaps 7;
363 LTVPRYKRDVOKILRQELSSQOQPQ--AGHCRIEVSREEIFEESYRQVMKRPDLWK 420
3617 LKAPRL-IDFDNKKAYFRSRIHQHCHISGPLRISVRRAYVLEDSYNQLMRSPQDLKG 3675
421 RLMIKFRGEGLDYGVAREWMLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
3676 RLNVQFQGEIGIDAGLTREWYQLLSRVIPDKGALLFTTVGNDATFQPNPNSVYQTEHL 3734
481 SYFHFVGRIMGVAFHGHYIDGGFTLPFYKQLLGKSIITDDMELVDPDLHNSLWILEND 540
3735 SYFKFVGRMVAKALFDGQLLDVYFTRSFYKHILGVKTVYHDIKAVDPDYKULWILEND 3794
541 ITGVLDHTFCVE-----HNAYGEIIQH 562
3795 VSDLDLTFSDMADEKHLIYKTEVRLMCFPCFFWCPIPKXCHICIELIILSLMKKVTDY 3854
563 ELKPNKKSIPVNEENKEYVRLVYNWRFURGIEAQFLALQKGFNEVIPHLLKTFDEKEL 622
3855 ELKPGGRNIRVTEETKHEYVDLVAGHILTNARIQINAFLEGFNELIPRELVSIFNDKEL 3914
623 ELIICGLKID-----VNDKWNTRLKHCTPDNSIVKFWKAVEFDEER 667
3915 ELLISGLPEIDCKLSDIQYLCAVPLDILKANTETSYTAGSEVIRFWFVWVKAFSKED 3974
668 RARLQFVTGSSRVPLOQFKALQGAAGPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPYES 727
3975 MARFQFVTGTSKVPLEGFKALQGISGQRLQIHKAYGAPERLPSAHCTFNQDLDFEYQS 4034
728 YEKLYEKLTLTAIETC-----GFA 746
4035 KEQLQERLLAIHASEGFGFA 4056
RESULT 14
S69625
hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C;Accession: S69625
R;Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A;Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A;Reference number: S69554
A;Accession: S69625
A;Molecule type: DNA
A;Residues: 1-3268 <DIE>
A;Cross-references: EMBL:U33050; NID:G927726; PIDN:AA564910.1; PID:G927738; MIPS:YDR45
C;Genetics:
A;Gene: SGD:TOM1
A;Cross-references: SGD:S0002865; MIPS:YDR457w
A;Map position: 4R
Query Match 19.5%; Score 789; DB 2; Length 3268;
Best Local Similarity 45.7%; Pred. No. 5.1e-48;
Matches 164; Conservative 59; Mismatches 132; Indels 4; Gaps 2;
384 SQOQQAQCHCRISREIPEESYRQVMKRPDLWKRLMIKFRGEGLDYGVAREWL 442
2905 NQREPK---LPITVRREQVFLDSYRALFFKTNDEIKNSKLEITFKGESGVDAAGVTREW 2961
443 YLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHLISYFHFVGRIMGVAFHGHYIDG 502
2962 QVLSQPMFNPDYALFLVPSDKITFHNRTSGINPEHLSFFKFIGMIGKAIKRDQCLDC 3021

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 protein - protein search, using sw model

on on: February 20, 2004, 15:25:26 ; Search time 35.0863 Seconds
(without alignments)
5501.382 Million cell updates/sec

File: US-10-009-945-4

Effect score: 4038

Sequence: 1 MSPFGRRRGPKVRLTVLC.....EKLYKLLTAIEETCGFAVE 748

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 830525 seqs, 258052604 residues

tal number of hits satisfying chosen parameters: 830525

imum DB seq length: 0

imum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

1: sp.archaea.*

2: sp.bacteria.*

3: sp.fungi.*

4: sp.human.*

5: sp.invertebrate.*

6: sp.mammal.*

7: sp.mhc.*

8: sp.organelle.*

9: sp.phage.*

10: sp.plant.*

11: sp.rodent.*

12: sp.virus.*

13: sp.veterebrate.*

14: sp.unclassified.*

15: sp.virus.*

16: sp.bacteriap.*

17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3041	75.3	728	11	Q8K300	Q8K300 mus musculus
2	2313	57.3	1061	5	Q9V853	Q9V853 drosophila
3	1910	47.3	376	13	Q98T84	Q98T84 xenopus lae
4	1688.5	41.3	335	11	Q8BSC0	Q8BSC0 mus musculus
5	1552	38.4	288	4	Q96DE7	Q96DE7 homo sapien
6	1418	35.1	786	3	O14326	O14326 schizosacch
7	1398	34.6	911	4	Q8NSA7	Q8NSA7 homo sapien
8	1386	34.3	955	4	Q96PU5	Q96PU5 homo sapien
9	1379	34.2	295	4	Q8ND8	Q8ND8 homo sapien
10	1378	34.1	238	11	Q9CSE3	Q9CSE3 mus musculus
11	1364.5	33.8	820	4	Q9NT88	Q9NT88 homo sapien
12	1358	33.6	887	11	Q8BGB3	Q8BGB3 mus musculus
13	1357.5	33.6	855	11	Q8CF10	Q8CF10 mus musculus
14	1355.5	33.6	887	11	Q62940	Q62940 rattus norv
15	1354.5	33.5	854	4	Q9HZM4	Q9HZM4 homo sapien
16	1354	33.5	835	11	Q9BRT9	Q9BRT9 mus musculus

043165 homo sapien

Q8WU9 homo sapien

Q9BW58 homo sapien

Q99PK2 mus musculus

O42573 xenopus lae

Q95R64 drosophila

Q54971 mus musculus

Q9Y0H4 drosophila

Q9BY75 homo sapien

O08758 mus musculus

Q96F66 homo sapien

O81QR5 drosophila

Q9H451 homo sapien

O43584 homo sapien

Q96CZ2 homo sapien

O00308 homo sapien

Q9H0M0 homo sapien

Q9VVI3 drosophila

Q9DBH0 mus musculus

O8BZ23 mus musculus

Q95XU3 caenorhabdi

Q9N2Z7 caenorhabdi

Q9BKW4 caenorhabdi

Q9STQ0 drosophila

O8T0C8 drosophila

O00307 homo sapien

Q9P2P5 homo sapien

Q9HCC7 homo sapien

Q9UTG2 schizosacch

ALIGNMENTS

RESULT 1

Q8K300 PRELIMINARY; PRT; 728 AA.

AC Q8K300; 01-OCT-2002 (TREMELrel. 22, Created)

DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)

DE Similar to E3 ubiquitin ligase SMURF1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

DR ENBL; BC029097; AAH29097.1; -.

DR InterPro; IPR000008; C2.

DR InterPro; IPR001202; WW_RSP5_WWP.

DR Pfam; PF00168; C2; 1.

DR Pfam; PF00832; HECT; 1.

DR Pfam; PF00397; WW; 2.

DR SMART; SM00239; C2; 1.

DR SMART; SM00119; HECTC; 1.

DR SMART; SM00456; WW; 2.

DR PROSITE; PS00499; C2 DOMAIN_1; 1.

DR PROSITE; PS00004; C2 DOMAIN_2; 1.

DR PROSITE; PS00237; HECT; 1.

DR PROSITE; PS01159; WW DOMAIN_1; 1.

DR PROSITE; PS00020; WW DOMAIN_2; 2.

DR KW Ligase.

SQ SEQUENCE 728 AA; 83083 MW; C355291B9D8AD757 CRC64;

Query Match 75.3%; Score 3041; DB 11; Length 728;

Best Local Similarity 75.0%; Pred. No. 2.5e-228;

Matches 572; Conservative 63; Mismatches 78; Indels 50; Gaps 7;

1 MSNPGRRNG-PVKRLRLTVLCAKLVKIDFFRLPDPFAKVVVDGSGQCHSTDTVKNLTDP 59
1 MSNPGTERRNGSSKIRLTVLCAKLVKIDFFRLPDPFAKVVVDGSGQCHSTDTVKNLTDP 60
60 KWOCHYDLYTGKSDSVTISVWNHKKIHKQAGFLGCVLLSNAINRLKDTGYQRLDLCK 119
61 KWSQHYDLYVGKSDSVTISVWNHKKIHKQAGFLGCVLLSNAINRLKDTGYQRLDLCK 120
120 LGPNNDNDTVRGQVWSLQSDRIGTGQVVDGSRFLDNDLPDGMERRTASGRQYLNHI 179
121 LNFSDTDAVAGQVWSLQSDRIGTGQVVDGSRFLDNDLPDGMERRTASGRQYLNHI 163
180 TRTQWERPRPASEYSSVSPRPLSCFVDENTPISGNGATCG-----QSSDPRLAIE 230
164 -----EDSGRPLSLCMEBPAPYDTGTGAAGCGNCRFVSPSDORLLV 209
231 RRVRSQHRNRYM-----SRTHLTPPLPGYEQRTTQGOVYFLHTQGVSTWHPVPR 286
210 QRLNPEVRGPIQTQNRPHGQSPPELPGYEQRTTQGOVYFLHTQGVSTWHPVPR 269
287 DLSNICEELGPIPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLNRQNLKD 346
270 DLNSVACDELGPPLPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLNRQNLKD 325
347 CQOQ-QVSLCPDDTECLVPRYKDLVOKLILRQELSQOQPOAGHCRIVSREEIFEE 405
326 PSQPLQPLSGSVDEBPAQRYERDLVOKLILRQELSQOQPOAGHCRIVSREEIFEE 385
406 SYRQVNMKPKDLMKLMKFRGEGLDYGVARWLYLLGHEMLNPFYGLFQVSRDIIY 465
386 SYRQIMKMRPKDKKELMYKFRGEGLDYGVARWLYLLGHEMLNPFYGLFQVSTNIIY 445
466 TLQINPDQSAVNEHLSYFFVORIMGAVFHGYTDGFTLPPFYKQLLGKSIITLDDMELV 525
446 TLQINPDQSAVNEHLSYFFVORIMGAVFHGYTDGFTLPPFYKQLLGKSIITLDDMELV 505
526 DPLHNSLVWILENDITGVLTDFCVRNAYGEIITQHELKPKNGKSIIPUNENKKEYRLY 585
506 DPLHNSLVWILENDITGVLTDFCVRNAYGEIITQHELKPKNGKSIIPUNENKKEYRLY 565
586 VNRWFLRGTEAFLAQKGFNVIPQHLKLTDFDEKLELIICGLGKIDVNDKVNTRLKH 645
566 VNRWFLRGTEAFLAQKGFNVIPQHLKLTDFDEKLELIICGLGKIDVNDKVNTRLKH 625
646 CTPDSNIVKWKAVFEFDEERARLLQVFGSSRVPLQGGKALQAGAPRLFTIHOIDA 705
626 CTVADSNVWFWQAVETFEERARLLQVFGSSRVPLQGGKALQAGAPRLFTIHOIDA 685
706 CTNNLPKATCFNRIDIPPEYSEYKLYEKLTLTAIETCGFAVE 748
686 NTDNLPKATCFNRIDIPPEYSEYKLYEKLTLTAIETCGFAVE 728
SULT 2
V853
QSV853 PRELIMINARY; PRT; 1061 AA.
QSV853; Q9U3W2;
01-MAY-2000 (trEMBLrel. 13, Created)
01-OCT-2002 (trEMBLrel. 22, Last sequence update)
01-MAR-2003 (trEMBLrel. 23, Last annotation update)
CG4943 protein (E3 ubiquitin ligase) (Ubiquitin-protein ligase) (Smad-ubiquitin E3 ligase Smuirl1).
LACK OR SMURF OR SMURF1 OR CG4943.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
STRAIN=Berkeley;
MEDLINE=20196066; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abriel J.P., Achavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D., Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Brannon R.C., Rogers Y., Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A., Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M., Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D., Ferrieria S., Frise E., Galle R.F., Garg N.S., George R.A., Gonzalez M., Houck J., Hoskins R.A., Hoston D., Howland T.J., Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A., McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J., Pacieb J., Paragas V., Park S., Patel S., Pfeiffer B., Phuanavanong S., Pittman G.S., Puri V., Richards S., Scheeler F., Stapleton M., Strong R., Svirskas R., Tector C., Tyler D., Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
"Sequencing of Drosophila melanogaster genome."
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K., Hradecky P., Huang Y., Kaminker J.S., Prochnick S.E., Smith C.D., Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E., Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N., Kronmiller B., Marshall B., Millburn G., Richter J., Russo S., Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
"Annotation of Drosophila melanogaster genome."
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
Flybase;
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```
[6]
SEQUENCE FROM N.A.
Laurencon A., Hawley S.;
"Molecular cloning of a type E3 Ubiquitin ligase.";
Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A.
Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;
"The Dsmurf ubiquitin-protein ligase restricts BMP signaling spatially
and temporally during Drosophila development.";
Dev. Cell 1:0-0 (2001).
[8]
SEQUENCE FROM N.A.
Liang Y.-Y., Lin X., Feng X.-H.;
"dSmurf1, a Smad-ubiquitin E3 ligase, specifically targets dpp-
activated Mad protein for degradation.";
Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
EMBL; AF003802; AAF57824.3; -
EMBL; AF216521; AAF21125.1; -
EMBL; AF416571; AAL09691.1; -
EMBL; AF464851; AAM09646.1; -
HSSP; Q13526; 1PIN.
FlyBase; Fgn0029006; lack.
InterPro; IPR000008; C2.
InterPro; IPR00569; HECT_domain.
InterPro; IPR002349; WW.
InterPro; IPR001202; WW_Rep5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00397; WW; 3.
PRINTS; PR00403; WWDOMAIN.
SMART; SM00239; C2; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2_DOMAIN_1; 1.
PROSITE; PS00499; C2_DOMAIN_2; 1.
PROSITE; PS50004; C2_DOMAIN_2; 1.
PROSITE; PS50237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 2.
PROSITE; PS50020; WW_DOMAIN_2; 3.
Ligase.
SEQUENCE 1061 AA; 115675 MW; 6BCC550F5129163 CRC64;
Query Match 57.3%; Score 2313; DB 5; Length 1061;
Best Local Similarity 45.1%; Pred.No.2.5e-171;
Matches 476; Conservative 107; Mismatches 157; Indels 316; Gaps 15;
7 RRNGVQLRLTVLCAKLVKDKFRLPDPFAKVVVGGSGCHSDTVKNTLDPKWNQHYD 66
8 RRNGTHKVRITLCARNLARKDLFRLPDPFAKVVVGGSGCHSDTVKNTLDPKWNQHYD 67
67 LYIGKSDSVTISVNHKKIKKQAGFLGCVRLLSNAINFLKDTGYQRLDCLKLPNDND 126
68 LFLGIGDAITVWNRQKHKH--GSGFLGCVRIAPFNIQSLKAGQFRLDLGKLSDDDE 125
127 TVRGIVVSLQSRD-----IGTGGVVDCSR--LFDNLPDGCWEERTASGRICQ 175
126 LVRGQIIISLSKDGPSGNPLAIVGSGDVRGSPEDSDSESLPGWEERTDGRVY 185
176 LNHITRTQWRPRTP----- 191
186 VNHTKSTQWRPRPQGVGVSSHATSPQORHNTNGSGDRQAPGTRSTCTNLMMNG 245
192 -----ASYSSGRLSCFV----- 206
246 HSRDLSTVADSRHSHTEILLSVGKENTFTFVSATTTTPGKTSNSSSSAGGRTLEQ 305
207 -----DENTPISGT-----N 216
306 RPTNEPAPTSTTSASVRLHSNDHVKTPKHQNGHAPSTPTSGTQQNVYNGAQN 365
217 GAT-----CGQSSDRLAER----- 231
366 GSTSGNGSGQAQPSASNGTWGTQDAATTSPSTTSPRHSQSPPTPNISPPATPSA 425
```

```

232 -----RVRSQRHNTMSRTHLTHPP----- 251
426 NGNVHSPNANSTPAGSGGSRSYAATPQGRSSRQQSGESSTRRSSRGTNGOTS 485
252 -----DLPEGYEORTTQQGVYFLHTQTGVSTWHPDR 283
486 GGGGGGGGORYASAAIAAQAARFPLDLPGYEMRTTQQGVYFHIPTGVSTWHPDR 545
284 VPRDL--SNINCELGPLPGWIRNATGRVYFDVHNNRTTQTDPRLSANLHLVLRQ 341
546 IPRDFDTQHLTDLAIGLPSGWEQRKTASGRVYFDVHNNRTTQTDPRLSGSLQIMRRG 605
342 N-----OLKQQQQQQVSL-----CPDDTE 361
506 TVPPTSANAGTAPPSPATPASAAPVPPQATPASNATPTLTITNPPHRIVDPDLPQ 665
362 CL-----TVPRYKRDVLQVKILRQELSOQQPQAGHCRIEVSREIIFESYQVMMRPK 416
666 GLLEGADLLPKYRDLVGKRLALRTELQTMQPSQGHCRLEVSREIIFESYRLIMKRAK 725
417 DLMKRLMIKRGEGGLDYGVAEWMYLLSHEMLYPYGLFOYSRDDIYTLQINPDSAVN 476
726 DMRKRLMVKFGEGLDYGVAEWMYLLSHEMLYPYGLFOYSRDDIYTLQINPDSAVN 785
477 PEHLSYFHFVGRIMGMAVFHGYIDGGFTLPFYKQLLKSITLDDMELVDPDLHNSLVI 536
786 PDHLSYFHFVGRITGIAVFHGHCLDGGFTTFFYKQLLKPITLGDIEGVDPLHRSLTWM 845
537 LENDITGLDHTFCVEHNAYGEIIQHELKPGKSPVNEENKCKEYVRLVYVWRFLRGIEA 596
846 LESNIGIIESTFVENNSFGALVWHELKPGASIPVTEENKREYVLYVYVWRFLRGIEQ 905
597 QFALQKGFNEVPOHLLKTFDEKELELIICGLKIDVNDKVNTRKHKCTPDSNIVKWF 656
906 QFALQKGFCELPISHLRPFDERELELVIGSISSIDVNDKVNTRKHKCTNETTQVLMF 965
657 WKAVEFPDEERRARLLQFVTGSSRVLPQGFKAQ---GAAGPRLFTIH-QIDACTNNLPK 712
966 WQVVEYSSEMRARLLQFVTGSSRVLPQGFKAQSTGAVGPRLETHLTADVPQTNLPK 1025
713 AHTCFNRIDIPPYESYKLVKLLTAIBETCGFAVE 748
1026 AHTCFNRIDIPPYETIQLLCKLQAVEITCGFAVE 1061
RESULT 3
Q98TS4
ID Q98TS4 PRELIMINARY; PRT; 376 AA.
AC Q98TS4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE E3 ubiquitin ligase Smurf2 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21107656; PubMed=11158580;
RA Zhang Y., Chang C., Gehling D.J., Hemmati-Brivanlou A., Derynck R.;
RT "Regulation of Smad degradation and activity by Smurf2, an E3
ubiquitin ligase.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:974-979 (2001).
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AY014181; AAG50422.1; -
DR HSSP; Q13526; 1PIN.
DR InterPro; IPR000008; C2.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00397; WW; 3.
DR SMART; SMC0239; C2; 1.
```

```
SMART: SM00456; WW; 3.
PROSITE; PS00004; C2_DOMAIN_2; 1.
PROSITE; PS01159; WW_DOMAIN_1; 1.
PROSITE; PS00020; WW_DOMAIN_2; 3.
Ligase.
NON TER
SEQUENCE 376 AA; 42652 MW; CA6AB2B5F4C20F98 CRC64;
Query Match
  47.3%; Score 1910; DB 13; Length 376;
  354; Conservative 10; Mismatches 8; Indels 4; Gaps 3;
1 MSNPGRRRANGPVKRLRLTVLCAKLVKKDFFR-LPDPFAKVVVDGSGQCHSTDTVKNTLDP 59
1 MSNQSRRNGPVKRLRLTVLCAKLVKKDFFGLPDSFAKVVVDGSGQCHSTDTVKNTLDP 60
60 KWNQHYDLYIGKSDSVISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGQRLDLCK 119
61 KWNQHYDLYIGKSDSVISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGQRLDLCK 120
120 LQPNNDIVRGQIVVVSQSRDRIGTGQGVVDCSRLFDNDLPDQWEERRTASGRIOVLNHI 179
121 LQPNNDIVRGQIVVVSQSRDRIGSGGVVDCSELFDNDLPDQWEERRTASGRIOVLNHI 180
180 TRTTOWERPTPASEYSPGRLSCFVDENTPIGTNGATCGQSSDPRLAERVRQRH 239
181 TRTTOWERPTPASEYSPGRLSCFVDENTPIGTNGASCGQTSDFRISERRVRQRH 240
240 NYMSTHLHTPPDLPEGVQRTTQGGVFLHTQTGVSTWHDPRVPRDLNINCELGPL 299
241 NYMSTHLHTPPDLPEGVQRTTQGGVFLHTQTGVSTWHDPRVPRDLNINCELGPL 300
300 PPGWEIRNTATGRVYFVDHNNRTQFTDPLSANLHLVLNRQN-QLKDDQQQQQVSLC-- 356
301 PPGWEIRNTATGRVYFVDHNNRTQFTDPLSANLHLVLNRQNQLKDDQQQQVSLCQL 360
357 PDTECLTVPRYKDL 372
361 PDEVECLTVPRYKDL 376
MULT 4
ISCO
Q8BSC0 PRELIMINARY; PRT; 355 AA.
Q8BSC0;
01-MAR-2003 (TRENBLrel. 23, Created)
01-MAR-2003 (TRENBLrel. 23, Last sequence update)
01-MAR-2003 (TRENBLrel. 23, Last annotation update)
Hypothetical HECT domain (Fragment).
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TIGSUS=Embryo;
MEDLINE=22354683; PubMed=12466851;
The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
Nature 420:563-573(2002).
EMBL; AK034736; BAC28813.1; -.
Hypothetical protein.
NON TER
SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;
Query Match
  41.3%; Score 1668.5; DB 11; Length 355;
  307; Conservative 25; Mismatches 20; Indels 3; Gaps 1;
397 VSREEIFESYRQIMKRPDKLKKLMVKFRGSEGLDYGGVAREWLYLLCHEMLNPPYGL 456
1 VSREEIFESYRQIMKRPDKLKKLMVKFRGSEGLDYGGVAREWLYLLCHEMLNPPYGL 60
457 FOYSRDDIYTLQINPDSAVNPEHLSYFPHVGRIMGMVAFHGHYIDGGFTLPFYKQLLGKS 516
61 FOYSTDNITYTLQINPDSINPDHLSYFPHVGRIMGLAVFHHYINGGFTVPFYKQLLGKP 120
517 ITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVEHNAYGRIIQLHKLKPNKSPVNEE 576
121 IQSDLESVDPELHKSLLWILENDITVLDHTFCVEHNAYGRIIQLHKLKPNRNPVTEE 180
577 NKKEYVRLYVNRFLRGIEAQLALQKGFNEVIPHLLKTFDEKELELIICGLGKIDVND 636
181 NKKEYVRLYVNRFLRGIEAQLALQKGFNELIPHLLKTFDEKELELIICGLDKIDLND 240
637 WKNTNLTCKTSDSNLVKMFKAVFEDEERARLLQFTGSSRVPLQGFALQ---CAA 693
241 WKNTNLTCKCADSNLVRFWQAVTFDEERARLLQFTGSTRVPLQGFALQGSTGAA 300
694 GPRLFTIHOIDACTNNLPKHAHTCFNRIDIPPYSEYKLYEKLKLTAEETCGFAVE 748
301 GPRLFTIHLIDANTDNLKHAHTCFNRIDIPPYSEYKLYEKLKLTAEETCGFAVE 355
RESULT 5
Q96DE7 PRELIMINARY; PRT; 288 AA.
AC Q96DE7;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Similar to E3 ubiquitin ligase SMURF2 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC009527; AA09527.1; -.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTc; 1.
DR PROSITE; PS00237; HECT; 1.
FT NON TER
SQ SEQUENCE 288 AA; 33255 MW; FE2B43E300E66537 CRC64;
Query Match
  38.4%; Score 1552; DB 4; Length 288;
  Best Local Similarity 100.0%; Pred. No. 8.4e-113;
  Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
461 RDDIYTLQINPDSAVNPEHLSYFPHVGRIMGMVAFHGHYIDGGFTLPFYKQLLGKSITLD 520
1 RDDIYTLQINPDSAVNPEHLSYFPHVGRIMGMVAFHGHYIDGGFTLPFYKQLLGKSITLD 60
521 DMELVDPDLHNSLVWILENDITGVLDHTFCVEHNAYGRIIQLHKLKPNKSPVNEE 580
61 DMELVDPDLHNSLVWILENDITGVLDHTFCVEHNAYGRIIQLHKLKPNKSPVNEE 120
581 YVRLYVNRFLRGIEAQLALQKGFNEVIPHLLKTFDEKELELIICGLGKIDVNDKVN 640
121 YVRLYVNRFLRGIEAQLALQKGFNEVIPHLLKTFDEKELELIICGLGKIDVNDKVN 180
641 TRLKCHTTPDSNIVKFWKAVFDEERRARLLQFTGSSRVPLQGFALQGAAGPRLFTI 700
181 TRLKCHTTPDSNIVKFWKAVFDEERRARLLQFTGSSRVPLQGFALQGAAGPRLFTI 240
701 HOIDACTNNLPKHAHTCFNRIDIPPYSEYKLYEKLKLTAEETCGFAVE 748
241 HOIDACTNNLPKHAHTCFNRIDIPPYSEYKLYEKLKLTAEETCGFAVE 288
```



```

Query Match          34.1% Score 1378; DB 11; Length 258;
Best Local Similarity 98.8%; Pred. No. 2.7e-99;
Matches 255; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

> 491 GMAVPHGYIDGGFTLPFYKQLGKSGITLDDMLVDPDLHNSLWILENDITGVLDHTFC 550
> 1 GMAVPHGYIDGGFTLPFYKQLGKSGITLDDMLVDPDLHNSLWILENDITGVLDHTFC 60
>
> 551 VERNAYGEIIQHELKPKNGKSI PVNEENKKEYVELYNWFLGIEAQFALOKGNEVIP 610
> 61 VERNAYGEIIQHELKPKNGKSI PVTEENKKEYVELYNWFLGIEAQFALOKGNEVIP 120
>
> 611 QHLKLTDEKELELIICGLKIDVNDKVNRLKCHCTPDSNIVKFWKAVEFFDEERAR 670
> 121 QHLKLTDEKELELIICGLKIDVNDKVNRLKCHCTPDSNIVKFWKAVEFFDEERAR 180
>
> 671 LLQFVTGSSRVLPQGFALOGAAGPRLFTTHQIDACTNNLPAHCTCFNRIDIPPYESYEK 730
> 181 LLQFVTGSSRVLPQGFALOGAAGPRLFTTHQIDACTNNLPAHCTCFNRIDIPPYESYEK 240
>
> 731 LYEXLLTAIBETCGFAVE 748
> 241 LYEXLLTAIBETCGFAVE 258
>

RESULT 11
INT88 PRELIMINARY; PRT; 820 AA.
> Q9NT88;
> 01-CCN-2000 (TrEMBLrel. 15, Created)
> 01-CCN-2000 (TrEMBLrel. 15, Last sequence update)
> 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
> Hypothetical protein (fragment).
> DKFZP342P2422.
> Homo sapiens (Human).
> Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
> Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
> [1]
> NCBI_TaxID=9606;
> SEQUENCE FROM N.A.
> TISSUE=Testis;
> Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
> Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
> -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
> EMBL; AL137469; CAB70754.1; -.
> HSSP; Q13526.1PIN.
> InterPro; IPR000008; C2.
> InterPro; IPR000569; HECT_domain.
> InterPro; IPR002349; WW.
> InterPro; IPR001202; WW_Rsp5_WWP.
> Pfam; PF00169; C2; 1.
> Pfam; PF00632; HECT; 1.
> Pfam; PF00397; WW; 3.
> PRINTS; PR00403; WWDOMAIN.
> SMART; SM00239; C2; 1.
> SMART; SM00119; HECTC; 1.
> SMART; SM00456; WW; 3.
> PROSITE; PS00004; C2_DOMAIN_2; 1.
> PROSITE; PS0237; HECT; 1.
> PROSITE; PS01159; WW_DOMAIN_1; 3.
> PROSITE; PS00020; WW_DOMAIN_2; 3.
> Hypothetical protein.
> NON_TER 1
> SEQUENCE 820 AA; 95283 MW; 0FD334B29B3F4123 CRC64;

Query Match          33.8% Score 1364.5; DB 4; Length 820;
Best Local Similarity 37.1%; Pred. No. 1.5e-97;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;

> 50 TDTVNTLDPKRNQHYLYIGKSD-SYTSVWNHKKHKQAGGFGVCR----- 98
> 12 TKTIKKTLNPKWNEEFYRVNPSNHLRLLFVFDENLRTDD---FLQVDVPLSHLPTD 68

```

```

QY 99 -----LLSNAINRLKDTGQYQBLDLCKLGPNDNDTVRGQIVVUSLOSDEIGTGG 146
DB 69 PTWERPYTKDFLLAPRSHKSKVKGFLRLKWAYPKNG-----QCDSENSDQRDDMEHGW 123
QY 147 QVVDSCRLFPND-----LPGWBEERTASGRIOYLNHITRTTOWERPT--- 189
DB 124 EVVD-----SNDASQHOBELPPLPPGPGWEEKVDNLGRTYVVAHNNRTTQWHRPSLMDV 178
QY 190 -----REASEYSSP-----CRPLSCFVDENTPISGYN-- 216
DB 179 SSESNDNIRQINQEAHRRFRSRRHISEDLEPSEGGDVPEPWET-ISEEVNIAGSLG 237
QY 217 -----GATCGSSDPR-LAB---RRVR-----SORHNTYMSRTHL----- 247
DB 238 LALPPPPASPGSKTSPQELSELSRLQITPDSNGEQFSSLIQREPPSSRLASCVTDAVA 297
QY 248 ---HTPPD-----LPEGYEORTTQGG 266
DB 298 EQGHLPPGAKDSFVRRAVKDTLSNFPQSPQSPYNSPKQHKVTSQSFPPGWMEMIAFNGR 357
QY 267 VYFLHTQTGVTWHDPRVP-----RDLNINCEBELGPLPGWEIRNTATGRVYFVDHNNR 321
DB 358 PFIDENTKTITTWEDPRLKFPVHVBKSTSLNPNDLGPLPGWEERIHLDCGTFYIDENSK 417
QY 322 TTQFTDPRLSANLHLVLRQNLKDQOQQQVVSLCPDDTECLTVPR--YKDLVQXKIL 379
DB 418 ITQWEDFRL-----QN-----PAITGPAVPYSGREFPKQKYDYF 449
QY 380 ROELSOQQPQAGHCRIEYSRESIPEESYRVQVMKMPKDLWK-RLMIKFRGEGLDYGVA 438
DB 450 RKLLKXPADIPNRFEMKLRHNNIFEESYRIRMSVKRPDVLKARLWIFESEKGLDYGVA 509
QY 439 REWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPSAV-NPEHLSYFHFVGRIMGMAYFHG 497
DB 510 REWFLLSKEMENPYGLFEYSATDNYTLQINPSGLCNEDHLSYFTFICRVAGLAVFHG 569
QY 498 HYIDGGFTLPFYKQLGKSGITLDDMLVDPDLHNSLWILENDITGVLDHTFCVHNAYG 557
DB 570 KLDDGFTIRFPYKMLGKQITLNDMESVDSYYSLKWILENDFT-ELDLMFCDIBENFG 628
QY 558 EIIQHELKPKNGKSI PVNEENKKEYVELYNWFLGIEAQFALOKGNEVIPQHLKTF 617
DB 629 QTVQVLDKNGSEIMVTNENKKEYIDLVIQWRFVNRVQKQNAFLEGFTTELLPDLIKIF 688
QY 618 DEKELELIICGLKIDVNDKVNRLK--HCTPDSNIVKFWKAVEFFDEERARILQFV 675
DB 689 DENELELLMCGLDVNDWRQHSIYKNGYC-PNHPVIQWFWKAVLLMDAEKRIQLQFV 747
QY 676 TGSSRVLPQGFALOGAAGPRLFTTHQIDACTNNLPAHCTCFNRIDIPPYESYEKLYEKL 735
DB 748 TGTSRVPMNGFAELYSGNGPQLFTIEQWS-PEKLPAHCTCFNRLDLPPTFPEDLREKL 806
QY 736 LTAIBETCGF 745
DB 807 LXAVENAQGF 816

RESULT 12
Q8BGE3 PRELIMINARY; PRT; 887 AA.
ID Q8BGE3
AC Q8BGE3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neural precursor cell expressed.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354693; PubMed=12466851;

```

[illegible]

```
SULT 14
2940
062940 PRELIMINARY; PRT; 887 AA.
01-NOV-1996 (Tremblrel. 01, Created)
01-NOV-1996 (Tremblrel. 01, Last sequence update)
01-MAR-2003 (Tremblrel. 23, Last annotation update)
Nedd4 (Fragment)
NEDD4.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
MEDLINE=96221297; PubMed=8665844;
Staub O., Dho S., Henry P., Correa J., Ishikawa T., McGlade J.,
Rotin D.;
"WW domains of Nedd4 bind to the proline-rich PY motifs in the
epithelial Na+ channel deleted in Liddle's syndrome.";
EMBO J. 15:2371-2380(1996).
-|- SIMILARITY: CONTAINS 1 C2 DOMAIN.
EVL; U50842; AAB48949.1; -.
HSP; Q13526; 1PIN.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 3.
PRINTS; PR00360; C2DOMAIN.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2 DOMAIN 1; 1.
PROSITE; PS00004; C2 DOMAIN 2; 1.
PROSITE; PS50237; HECT; 1.
PROSITE; PS01159; WW DOMAIN 1; 3.
PROSITE; PS50020; WW DOMAIN 2; 3.
NON_TER 887 887
SEQUENCE 887 AA; D74B1097688CD9A1 CRC64;

Query Match
Best Local Similarity 37.1%; Score 1355.5; DB 11; Length 887;
Matches 314; Conservative 131; Mismatches 248; Indels 153; Gaps 25;

14 LRTVLCAKLVKDFRLPDPFAKV-----VDGSGQCHSTDTKNTLDPKWNQHYDLY- 68
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
77 VRVKVLAGLAKKQILGASDFVVRVTLVDPMSGVLTSVQTKIKSLNPKWNEI-LFR 135
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 -ICKSDSVISVWVHKKIHKQAGFLGVR-----LJLSNAIN 105
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 VLPQQHRIIEFVFNRLFRD---FLGGVDVPLVPLTENPRMERPYTFKDFVLHPRSH 192
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 RLKXDTGYRLDLCKLGNPDNDTVRGQIVSLQSRDRIGTGQVVD---CSRL-----FDN 157
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193 KSRVGYALKVLYLPKNGSD-----ENADQAELEFGVWVLDQDPAATHLOHPPEPS 246
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
158 DLPDGHREARTASGRQYLNHTRTQWERPT-----RPAS 193
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 PLPPGWEERQDVLGRYYVNHESRTOMKRPSPEDDLTDENGDIQQAAGFTTRQIS 306
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 E-----YSSPORPLSCFVDNTPISTNGATCGSSDP-----RLAE----- 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
307 EDVDPGDNHESPENWEIVREDNTIYSG-QAVQSPSPGHPDVQVRLAEELDTRLTWVGNP 365
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
231 ---RRVRSQH---RNTVST-----HLHTPDLPEGYEQRTQQQGVYFLHTQT 274
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
366 ATSPQVTSNHSRSGSSQTSFEEQPTLPVLLPTSSGLPPGWEKQDGRSRYVDHNS 425
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
275 GVSTWHPD-----RVP---RDLNSINCELGPLPGWEIRATGATGVYFVDHNRRT 323
```

```
Db 426 KTTWSKPTMQDDPSKIPAHLRGKTPVDSNDLGLPFGWEERTHTDGRVFFINHIKKT 485
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 QFTDRLSANLHLVLRNQQLKDOQQQVSLCPDDTECLTPRYKEDLVQKLKILRQEL 383
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 486 QWEDPRM-----QNVAITGPAEP-----YSRDYKRYEYFFRKL 519
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 SOQQAQAGHCRIEVSREEIFESYRQVWKVPEKDLWK-RLMIKPRGEGGLDYGVAREWL 442
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 520 KQTDIPNKFEMKLRANILEDSTYRIMGVVRADFLKARLWIEFDGEGGLDYGVAREWF 579
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 443 YLLSHEMLNPYGLFQYS-RDDIYTLQINPDSAV-NPEHLSYHFVGRIMGMAVFFGHYI 500
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 580 FLISKEMFNPYGLFEYSATEDNTVLIQNPNSGLCNEDHLSYFKFGRVAGMAVYHGKLL 639
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 501 DGGFTLPYVKOLLGKSITLDDMELVDPDLNLSLWILENDITGVLDHTFCVEHNAYGEII 560
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 640 DGGFTRPYKMWLQKLITLHDWESVDSEYSSRWILENDPT-ELDLRFIDBELFGQTH 698
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 561 QHELPNGKSIPIVNEENKKEYVRLVYVNRFLRGIEAQFLALQKGFNEVIPOHLLKTFDEK 620
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 699 QHELKTGGSEVVVTVNKNKKEYIYLVIQWRFVNRIOQMFAKSGFFELIPQDLIKIFDEN 758
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 621 ELELIICGLKIDVNDKVNTRLKH-CTPDSNIVKWPWKAVERFDEBERRARILQFTVGSS 679
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 759 ELELLMCGLDVNDVNDREHTKYKNGYSLNHQVHWFVKAVLWMDSEKIRLLQFVTGTS 818
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 680 RYPLQGFKALQAGAPRLFTIHQIDACTNNLPKACTCNFRIDIPPYESYEKLYEKLITAI 739
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 819 RYPMNGFAELYGSNGPSQFTVEQW-GTFDKLPRAHTCFNRLDIPPYESFDELWKLQMAI 877
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 740 BETQGF 745
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 878 ENTQGF 883
```

RESULT 15

```
Q9H2W4
ID Q9H2W4 PRELIMINARY; PRT; 854 AA.
AC Q9H2W4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE NEDD4La (Ubiquitin ligase NEDD4La).
GN NEDD4La OR NEDD4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Chen H., McInnis M.G., DePaulo R. Jr., Ross C.A.;
RT "Identification of a novel gene on human chromosome 18q21 with
RT homology to NEDD4."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Chen H., Ross C.A., Wang N., DePaulo R. Jr., McInnis M.G.;
RT "NEDD4L on human chromosome 18q21 has multiple forms of transcripts
RT and is a novel member of the NEDD4 gene family."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210730; AAG43524.1; -.
DR EMBL; AF385931; AAM46208.1; -.
DR HSP; Q13526; 1PIN.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WNDOMAIN.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 4.
```

```

Ligase.
SEQUENCE      854 AA;  98181 MW;  00C74E161F52E7F CRC64;

Query Match      33.5%;  Score 1354.5;  DB 4;  Length 854;
Best Local Similarity 43.7%;  Pred. No. 9.9e-97;
Matches 280;  Conservative 98;  Mismatches 200;  Indels 63;  Gaps 13;

137 QSRDRIGTGQGVVDCSRUFDN---DLPGWEERTASGRIOYLNHIHITRTQWERPTRPAS 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 RARSSTVTGGEPTSAVAVHTTGPLSGWEERKDAKRTYYVNNHNRRTTWTTRPIMQLA 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

194 EYSSPG-----RPLSC---FVDENTPISGINGATCGOSSDPRLAERVRVS 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 EDGASGSAIANNHLEIEPQIRPRSLSPVTTLGAPLEGAKDSPVRAVXDTLSNPSPQ 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

236 QHRNVMSTHLHTPDLPEGYEQRTTQGGVYFLHTGTGVSTWHDPRVP-----RDLN 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 PSFYNPKFKHKVQTSFUPPGWEMRIANGRPFFIDHNTKTTIWTWEDPRLKFPVHMRSKTS 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

291 INCEELGPLPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLNRQNLKDOQQQ 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
421 LNENDLGPLPPGWEIRIHLDTGTFYIDENSKITQWEDPRL-----QN----- 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

351 QVVSICPDDETECLTVPR--YKDLVQKLIKRLQELSQOQPOAGHCRIEVSREBIFESYR 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
463 -----PAITGPVAVPYSEFKQKYDYFRKKLKPADIPNRFENFKLHNNIFEESYR 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

409 QVMQMRPKDLWK-RLMIKFRGEGLDYGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTL 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
513 RIMSVKRPDVLKARLWIEFESKGLDYGVAREWFFLLSKEMFNFPYGLFEYSATDNYTL 572
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

468 QINPDSAV-NPEHLSYTHFVGRIMGMAVFGHYIDGGFTLFPYKQLLGKSIITLDDMELVD 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
573 QINPNSGLCNEHLSYTHFVGRVAGLAVFHGKLLDGGFFIRFPYKMLGKQITLNDMESVD 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

527 PDLHNSLWILENDITGLDHTFCVHEHNAIGEIIQHELKPNKSIIPYNEENKKEYVLYV 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
633 SEFYNSLWILENDPT-ELDLNFCIDENFGCTQYQVDLKPNGSEIMVTNENKREYIDLVI 691
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

587 NNRFLRGIEAQLAQGFNEVIPQHLKTFDEKELELIICGLGKIDVNDKVNTRLK-- 644
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
692 QNRFVNRVQKQMAFLEGFTLLPIDLIKIFDENELELLMCGLDGVDVNDWRQHSIYKNG 751
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

645 HCTPDSNIYKMWKAVEFDEERRARILQFVTGSSRVPLOGFKALQAGAPRLFTIHQID 704
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
752 YC-FNHPVIQWFVKAVLLMDAEKRIILQFVTGTSRVPVPMNGFAELYGNGPQLFTIEQWG 810
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

705 ACTNNLPKHAHTCFNRIDIPPYESYEKLYEKLITAEETCGF 745
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
811 S-PEKLPRAHTCFNRLDLPPEYTFEDLREKULMAVENAQGF 850
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

arch completed: February 20, 2004, 15:29:57
b time : 39.0863 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: February 20, 2004, 15:25:26 ; Search time 11.1869 Seconds
(without alignments)
3144.377 Million cell updates/sec

itle: US-10-009-945-4
effect score: 4038
equences: 1 MSNPGRRNGPVKRLTVLC.....EKLVEKLLTAIBETCGPAVE 748

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 127863 seqs, 47026705 residues

total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	4031	99.8	748	1 SUP2_HUMAN	Q9hau4 homo sapien
2	3015.5	74.7	757	1 SUP1_HUMAN	Q9hce7 homo sapien
3	2937.5	72.7	731	1 SUP1_XENLA	Q9pun2 xenopus lae
4	2491.5	61.7	619	1 SUP1_MOUSE	Q9cun6 mus musculus
5	1585.5	39.3	767	1 PUB1_SCHPO	Q92462 schizosacch
6	1484	36.8	809	1 RSP5_YEAST	P39940 saccharomyc
7	1358	33.6	957	1 NED4_MOUSE	P46935 mus musculus
8	1354.5	33.5	927	1 NED4_HUMAN	P46934 homo sapien
9	626	15.5	310	1 URB1_FAT	P51593 rattus norv
10	539	13.3	875	1 UR3A_HUMAN	Q05086 homo sapien
11	484.5	12.0	885	1 UR3A_MOUSE	Q08759 mus musculus
12	460.5	11.4	892	1 HUL4_YEAST	P40985 saccharomyc
13	445.5	11.0	1050	1 HER3_HUMAN	Q15034 homo sapien
14	404	10.0	1092	1 TRIB_HUMAN	Q14669 homo sapien
15	342	8.5	910	1 HUL5_YEAST	P53119 saccharomyc
16	286.5	7.1	1620	1 HED1_HUMAN	Q9ult8 homo sapien
17	281	7.0	920	1 EDD_FAT	Q82671 rattus norv
18	280	6.9	2799	1 EDD_HUMAN	Q95071 homo sapien
19	278.5	6.9	1483	1 YF4_YEAST	P33202 saccharomyc
20	262.5	6.5	1647	1 YDE1_SCHPO	Q10435 schizosacch
21	224.5	5.6	2895	1 HYD_DROME	P51592 drosophila
22	207	5.1	472	1 YAP1_MOUSE	P46938 mus musculus
23	158	3.9	448	1 YAP1_CHICK	P46936 gallus gall
24	154.5	3.8	454	1 YAP1_HUMAN	P46937 homo sapien
25	122.5	3.0	672	1 KPFA_RABIT	P17252 homo sapien
26	122.5	3.0	672	1 KPFA_RAT	P10102 oryctolagus
27	122.5	3.0	672	1 KPFA_MOUSE	P05696 rattus norv
28	122	3.0	637	1 GVP1_YEAST	Q08484 saccharomyc
29	120.5	3.0	1813	1 UN13_CABEL	P27715 caenorhabdi
30	119.5	3.0	593	1 CNE5_HUMAN	Q9hch3 homo sapien
31	116.5	2.9	583	1 P440_YEAST	P33203 saccharomyc
32	116.5	2.9	672	1 KPFA_MOUSE	P20444 mus musculus
33	116	2.9	163	1 PIN1_HUMAN	Q13526 homo sapien

ALIGNMENTS

RESULT 1

ID	SUP2_HUMAN	STANDARD;	PRT;	748 AA.
AC	Q9HAU4: Q9H260;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DE	Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--			
DE	protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)			
DE	(hSMURF2).			
GN	SMURF2			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-251--VAL-284 AND			
RP	GLY-297--LEU-330.			
RX	PubMed=1163210;			
RA	Kavsek P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,			
RA	Thomsen G.H., Wrana J.L.;			
RT	"Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the			
RT	TGF-beta receptor for degradation."			
RL	Mol. Cell 6:1365-1375(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-251--VAL-284 AND			
RX	MEDLINE=20538422; PubMed=11016919;			
RA	Lin X., Liang M., Feng X.-H.;			
RT	"Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent			
RT	degradation of Smad2 in transforming growth factor-beta signaling."			
RL	J. Biol. Chem. 275:36818-36822(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF CVS-716.			
RX	MEDLINE=21107656; PubMed=1158580;			
RA	Zhang Y., Chang C., Gehling D.J., Hemmati-Briavanlou A., Derynck R.;			
RT	"Regulation of Smad degradation and activity by Smurf2, an E3			
RT	ubiquitin ligase."			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).			
CC	-!- FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to			
CC	trigger their ubiquitination and proteasome-dependent degradation.			
CC	Enhances the inhibitory activity of SMAD7 and reduces the			
CC	transcriptional activity of SMAD2. Coexpression of SMURF2 with			
CC	SMAD1 results in considerable decrease in steady-state level of			
CC	SMAD1 protein and a smaller decrease of SMAD2 level.			
CC	-!- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but			
CC	not SMAD4.			
CC	-!- SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of			
CC	SMAD7.			
CC	-!- TISSUE SPECIFICITY: Widely expressed.			
CC	-!- DOMAIN: The second and third WW domains are responsible for			
CC	interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).			
CC	-!- SIMILARITY: Contains 1 C2 domain.			
CC	-!- SIMILARITY: Contains 3 WW domains.			
CC	-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase			
CC	domain.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
EMBL; AF310676; AAG45422.1; -
EMBL; AF310143; AAG25641.1; -
EMBL; AY014180; AAG50421.1; -
HSP; Q13526; 1PIN.
MIM; 605532; -.
GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.
GO; GO:0016481; P:negative regulation of transcription; NAS.
GO; GO:0017015; P:regulation of TGFbeta receptor signaling pa...; NAS.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 3.
SMART; SM00239; C2; 1.
SMART; SM00119; HECT; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2 DOMAIN_1; 1.
PROSITE; PS00004; C2 DOMAIN_2; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW DOMAIN_1; 1.
PROSITE; PS00020; WW DOMAIN_2; 3.
Ubl conjugation pathway; Ligase; Repeat; Nuclear protein.
DOMAIN 1 98 C2 DOMAIN.
DOMAIN 157 WW 1.
DOMAIN 251 WW 2.
DOMAIN 297 WW 3.
DOMAIN 414 HECT.
MUTAGEN 251 284
MUTAGEN 297 330
MUTAGEN 716
MUTAGEN 716
CONFLICT 6 6
SEQUENCE 748 AA; 86195 MW; 3042B443A3755762 CRC64;

Query Match          99.8%; Score 4031; DB 1; Length 748;
Best Local Similarity 99.9%; Pred. No. 1.6e-280;
Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSNPGRRNGVKRLTLVCAKNLVKDFFRLPDPFAKVVVGSGGCHTDVQNTLDPK 60
2 1 MSNPGRRNGVKRLTLVCAKNLVKDFFRLPDPFAKVVVGSGGCHTDVQNTLDPK 60
61 WQHYDLVIGKSDSVTSVWNHKKHKKQAGLGCVRLLSNAINLKDGTGYRLDCLK 120
61 WQHYDLVIGKSDSVTSVWNHKKHKKQAGLGCVRLLSNAINLKDGTGYRLDCLK 120
121 GPNDNDTVRGQIVVVSLOSRRIGTGQGVDCSLFNDLPDQWEERTASGRIQYLNHIT 180
121 GPNDNDTVRGQIVVVSLOSRRIGTGQGVDCSLFNDLPDQWEERTASGRIQYLNHIT 180
181 RTTQWERPTRASSYSPGRLSCFVDENPTISGTGATCGSSDPRLAERVRQRHNRN 240
181 RTTQWERPTRASSYSPGRLSCFVDENPTISGTGATCGSSDPRLAERVRQRHNRN 240
241 YMSRTHLTPDLPDPEGYEQRITQGGQVYFLHTGTGSTVHDPVRPDLNSINCEELGFLP 300
241 YMSRTHLTPDLPDPEGYEQRITQGGQVYFLHTGTGSTVHDPVRPDLNSINCEELGFLP 300
```

```
Qy 301 PGWEINTATGRVYFDVHNRTTQFTDPRLSANLHLVLRQNLKDQOQOQVWSLCFDDT 360
Db 301 PGWEINTATGRVYFDVHNRTTQFTDPRLSANLHLVLRQNLKDQOQOQVWSLCFDDT 360
Qy 361 ECLTVPRYRDVLQKILRLQELSSQOQAGHCRLEVSREEIFFEESYRQVMKMPKOLWK 420
Db 361 ECLTVPRYRDVLQKILRLQELSSQOQAGHCRLEVSREEIFFEESYRQVMKMPKOLWK 420
Qy 421 RLMIKFRGEGLDYGVAVREWLILLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db 421 RLMIKFRGEGLDYGVAVREWLILLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Qy 481 SYFHFVGRIMGVAIPHGVIDGGFTLPVFKQLLGSITLDDMELVDPDLHNSLVWILEND 540
Db 481 SYFHFVGRIMGVAIPHGVIDGGFTLPVFKQLLGSITLDDMELVDPDLHNSLVWILEND 540
Qy 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFJRGIEAQFLA 600
Db 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFJRGIEAQFLA 600
Qy 601 LQKGFNEVIPHLLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Db 601 LQKGFNEVIPHLLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Qy 661 EFDEERRARLLQFVTGSSRVLQGFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 720
Db 661 EFDEERRARLLQFVTGSSRVLQGFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 720
Qy 721 DIPVESYKLYEKLITAEETCGFAVE 748
Db 721 DIPVESYKLYEKLITAEETCGFAVE 748

RESULT 2
SUFI_HUMAN STANDARD; PRT; 757 AA.
AC QHCE7; O75853; Q9UUT8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
DE (hSMURF1).
GN SMURF1 OR KIAA1625.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).
RX MEDLINE=99385348; PubMed=10458166;
RT Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RT pattern formation.";
RL Nature 400:687-693(1999).
RN [2]
RP SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
RA Stoneking T., Bauer C., O'Neal D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII: the complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
CC -1- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION.
CC -1- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
```



```

Name=Long;
Isid=Q9HCE7-1; Sequence=Displayed;
Name=Short;
Isid=Q9HCE7-2; Sequence=VSP_006812;
-|- SIMILARITY: Contains 1 C2 domain.
-|- SIMILARITY: Contains 2 WW domains.
-|- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF199364; AAF08298.2; -
EMBL; AC004893; AAC62434.1; -
EMBL; AB346845; BAB13451.1; ALT_INIT.
HSP; Q13526; IPIN.
MIN; 605568; -
GO; GO:0005622; C:intracellular; TAS.
GO; GO:0000211; F:protein degradation tagging activity; IDA.
GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO; GO:0030154; P:cell differentiation; IDA.
GO; GO:0007398; P:ectoderm development; TAS.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
GO; GO:0006464; P:protein modification; TAS.
GO; GO:0006512; P:ubiquitin cycle; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR01202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00499; C2 DOMAIN 1; 1.
PROSITE; PS00004; C2 DOMAIN 2; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW DOMAIN 1; 1.
PROSITE; PS0020; WW DOMAIN 2; 2.
UBI conjugation pathway; Ligase; Repeat; Alternative splicing.
DOMAIN 1 99 C2 DOMAIN.
DOMAIN 234 267 WW 1.
DOMAIN 306 339 WW 2.
DOMAIN 420 757 HECT.
BINDING 725 725 UBIQUITIN.
VARSPIC 269 294 Missing (in isoform Short).
/FTIdVSP_006812.
MUTAGEN 725 725 C->A; LOSS OF UBIQUITINATION CAPACITY.
SEQUENCE 757 AA; 86113 MW; 89A171CFC47B40E9 CRC64;

Query Match
Best Local Similarity 74.7%; Score 3015.5; DB 1; Length 757;
Matches 574; Conservative 63; Mismatches 73; Indels 85; Gaps 10;

1 MSNPGRRNG-PVKRLTLVLCALVKKDPFPLPDPFAKVVDGSGQCHSDTVKNTLDP 59
1 MSNPGTRNGSSIKRLTLVLCALVKKDPFPLPDPFAKVVDGSGQCHSDTVKNTLDP 60
60 KWNQHYDLVIGKSDSVTISVWNHKKHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCK 119
61 KWNQHYDLVIGKSDSVTISVWNHKKHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCK 120
120 LGNDNDTVRGQIVWSLQSDRIGTGQGVQVDCSLFNDLDPDGEERTASGRQYLNHI 179
121 LNPSTDVARGQIVWSLQSDRIGTGQGVQVDCSLFNDLDPDGEERTASGRQYLNHI 163
180 TRTOWERPTRPASEYSSGRPLSCFVDENTPISGTNGATCG-----QSSDPLAE 230

```

```

Db 164 -----EDSGRPLSCFMEEPAPYTDSTGAAAGGNCRFVSPSQDQLQA 209
QY 231 RVRVSQRHNYM-----SRTHLTPDPLPEGVEQTTQGGQVYFLHTOTGVSTWHDPRV-- 284
Db 210 QRLRNPDRVGRSLQTPQNRPHGQSPFELPEGVEQTTQGGQVYFLHTOTGVSTWHDPRIPS 269
QY 285 -----PRDLSNINCEELGFLPGWEIRNATGRVYFVDHNN 320
Db 270 PSGTIPGGDAAPLYBFLQCHTSFRLNSVNCDELGFLPGWEVRSTVSGRIYFVDHNN 329
QY 321 RTQTDRPLSNHLNLRNQKLDKQKQV-----SLCPDDTECLTVPRYKRDVLVKL 376
Db 330 RTQTDRP-----LHIMHQCLKEPSQPLPSEGL-----EDELPQAKYERDLVKL 382
QY 377 KILROELSQOQAGCHRIEVSREIEPESVROVMKRPKDLKRLMIKRGEGGLDYG 436
Db 383 KVLRELSLQOQAGCHRIEVSREIEPESVROVMKRPKDLKRLMIKRGEGGLDYG 442
QY 437 VAREWLYLLSHEMLNPPYGLFYGRDDIYTLQINPDSAVNPEHLSYFHFVGRIMGMAVFH 496
Db 443 VAREWLYLLCHEMLNPPYGLFYGRDDIYTLQINPDSAVNPEHLSYFHFVGRIMGMAVFH 502
QY 497 GHVIDGGFTLPEYKOLLGKSLTLDMLVDPDLNLSVLENDITGVLDHTFCVENAY 556
Db 503 GHYINGGFTVPYKOLLGKSLTLDMLVDPDLNLSVLENDITGVLDHTFCVENAY 562
QY 557 GRIHQELKPKNGKSIPIVNEENKKEVRLYVNWRLRGIEAQFALQKGFNEVIPHLLKT 616
Db 563 GRIHQELKPKNGKSIPIVNEENKKEVRLYVNWRLRGIEAQFALQKGFNEVIPHLLKT 622
QY 617 FDEKELELIICGLKIDVNDKVNTRLKCHTTPSDNIVKWFKAVEFDEERRARLLQFVT 676
Db 623 FDEKELELIICGLKIDVNDKVNTRLKCHTTPSDNIVKWFKAVEFDEERRARLLQFVT 682
QY 677 GSSRVPLQGFALQ-----GAAGFRLTIHQDACTNNLPKATCNRIDIPYSEYKLYE 733
Db 683 GSTRVPLQGFALQSGTGAAGFRLTIHQDACTNNLPKATCNRIDIPYSEYKLYE 742
QY 734 KLLTAIBETCGFAVE 748
Db 743 KLLTAIBETCGFAVE 757

RESULT 3
ID SCFL_XENLA STANDARD; PRT; 731 AA.
AC Q9PUN2;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (SC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase) (xSMURF1).
GN SMURF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blastula;
RX MEDLINE=99385348; PubMed=10459166;
RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RL pattern formation."
RL Nature 400:687-693(1999).
CC -|- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
CC REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
CC SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2
CC (ACTIVIN/TGF-BETA) PATHWAY.
CC -|- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING

```

TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC ECTODERM AND INVOLVING MESODERM, AND EXPRESSION GRADUALLY LOCALIZES TO THE NERVOUS SYSTEM. AT EARLY TADPOLE STAGES EXPRESSED IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.

!- SIMILARITY: Contains 1 C2 domain.

!- SIMILARITY: Contains 2 WW domains.

!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF169310; AAD52564.1; ..
 GO; GO:0005622; C:intracellular; TAS.
 GO; GO:000211; F:protein degradation tagging activity; IDA.
 GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
 GO; GO:0030154; P:cell differentiation; IDA.
 GO; GO:0007398; P:ectoderm development; TAS.
 GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
 GO; GO:0006464; P:protein modification; TAS.
 GO; GO:0006512; P:ubiquitin cycle; IDA.
 InterPro: IPR000008; C2.
 InterPro: IPR000569; HECT domain.
 InterPro: IPR001202; WW_Rsp5_WWP.
 Pfam; PF00168; C2; 1.
 Pfam; PF00632; HECT; 1.
 Pfam; PF00397; WW; 2.
 SMART; SM00239; C2; 1.
 SMART; SM00119; HECT; 1.
 SMART; SM00456; WW; 2.
 PROSITE; PS00499; C2_DOMAIN_1; 1.
 PROSITE; PS00004; C2_DOMAIN_2; 1.
 PROSITE; PS0237; HECT; 1.
 PROSITE; PS01159; WW_DOMAIN_1; 1.
 PROSITE; PS0020; WW_DOMAIN_2; 2.
 Ubl conjugation pathway; Ligase; Repeat.
 DOMAIN 1 99 C2 DOMAIN.
 DOMAIN 233 266 WW 1.
 DOMAIN 279 312 WW 2.
 DOMAIN 394 731 HECT.
 SEQUENCE 731 AA; 83259 MW; 3CB885512A42CE2C CRC64;

Query Match 72.7%; Score 2937.5; DB 1; Length 731;
 Best Local Similarity 72.7%; Pred. No. 2.5e-202;
 Matches 559; Conservative 68; Mismatches 85; Indels 57; Gaps 9;

1 MSNPGRRNG-PVKLELTLCANLVKDFRLPPFAKVVDSGQCHSTDTVNTLDP 59
 GO: GO:0005622; C:intracellular; TAS.
 1 MSNVVTRGGSSIRVTLVLCANLAKEDFFRLPPFAKVVDSGQCHSTDTVNTLDP 60
 GO: GO:0005622; C:intracellular; TAS.
 60 KMNQHYDLYIGSDSVTSVWNNKHKKQAGFLGCVRLLSNAINRLKDTGYORLDLCK 119
 61 KMNQHYDLYIGSDSVTSVWNNKHKKQAGFLGCVRLLSNAINRLKDTGYORLDLCK 120
 120 LGPNDNDVTRGQIVVLSQDRIGTGQVVDGSRFLDNDLDPGWERTAGRIQYLNHI 179
 121 LNPTDNDVTRGQIVVLSQDRIGTGQVVDGSRFLDNDLDPGWERTAGRIQYLNHI 162
 180 TRTTQWERPRPASEYSSRPLSCPDVENTIS---GNGATCGSSDPLAERRVRSQ 236
 163 -----LEDTGGRPLSCPDVENTIS---GNGATCGSSDPLAERRVRSQ 209
 237 RHRNYNSRTHLTP-----PDLPEGYEORTTQQGVYPLHTQTGVSTWHPDPRVPRD 287
 210 RVRGPEVREHVQTPQNRSHFGSQDLPEGYEORTTQQGVYPLHTQTGVSTWHPDPRD 269
 288 LSNINCELGPLPGWEIRNTATGRVYFVDHNNRTTQTDPRLSNLHLVLRNQNLKQ 347

Db 270 LNSVNCDDLSLSPAGNEVTRTVSGRIYFVDHNNRTTQFTDPR-----LHHIHHQSLKEP 325
 QY 348 QQ-----QQVSLQDDTECLTVPRYKRDVLQKILRQELSSQQQPOAGHCRVEVSREEIF 403
 Db 326 NHAIFVQSDGSL--EDGDEFFAQRYSRDLVQKLVLRHELSSLQPOAGHCRVEVSREEIF 383
 QY 404 BESRYQVMKRPKDLWKRLMKFERGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDD 463
 Db 394 BESRYQVMKRPKDLWKRLMKFERGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSTDN 443
 QY 464 IYTLQINPDSAVNPEHLSYFHFVGRINGMAVFGHVIDGGFTLPFYKQLLGKSLITLDNE 523
 Db 444 IYTLQINPDSINPDHLSYFHFVGRINGMAVFGHVIDGGFTLPFYKQLLGKSLITLDNE 503
 QY 524 LVDPDLNSLWILENDITGVLDHTFCVHNAYGEIIOHELKPNKSIIPVNEENKKEYVR 583
 Db 504 SVDFELKSLWILENDITGVLDHTFCVHNAYGEIIOHELKPNKSIIPVNEENKKEYVR 563
 QY 584 LYVNRFLRGIEAQLALQKGFNEVPIQHLKTFDEKLELICGLKIDVNDKVNTEL 643
 Db 564 LYVNRFLRGIEAQLALQKGFNEVPIQHLKTFDEKLELICGLKIDVNDKVNTEL 623
 QY 644 KHCTPDSNIVKFWKAVEFEDEERRALLQFVTGSSRVPLQGFALQ---GAAGPRLFTI 700
 Db 624 KHLANSNIVQWFWQAVESFEDEERRALLQFVTGSSRVPLQGFALQ---GAAGPRLFTI 683
 QY 701 HQIDACTNNLPKACHTCFNRIDIPPEYSEKLYEKLTAIBETCGFAVE 748
 Db 684 HLIDANTDNLKPAHTCFNRIDIPPEYSEKLYEKLTAIBETCGFAVE 731

RESULT 4

SUFI_MOUSE
 ID SUFI_MOUSE STANDARD; PRT; 619 AA.
 AC Q9CUN6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
 protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
 GN (Fragment).
 DE SMURF1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC !- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
 BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR

UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION (By similarity). Contains 1 C2 domain.
-!- SIMILARITY: Contains 2 WW domains.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AK015264; BAB29770.1; --
HSP; Q13526; IPIN
MGD; MG11923038; 4930431E10Rik.
GO: GO:0005622; C:intracellular; ISS.
GO: GO:0002111; F:protein degradation tagging activity; ISS.
GO: GO:0004842; F:ubiquitin-protein ligase activity; ISS.
GO: GO:0030154; P:cell differentiation; ISS.
GO: GO:0007398; P:ectoderm development; ISS.
GO: GO:0030514; P:negative regulation of BMP signaling pathway; ISS.
GO: GO:0006464; P:protein modification; ISS.
GO: GO:0006512; P:ubiquitin cycle; ISS.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00499; C2_DOMAIN_1; PARTIAL.
PROSITE; PS00004; C2_DOMAIN_2; PARTIAL.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 1.
PROSITE; PS00020; WW_DOMAIN_2; 2.
Ubl conjugation pathway; Ligase; Repeat.
NON_TER 1
DOMAIN 125 158 WW 1.
DOMAIN 171 204 WW 2.
DOMAIN 285 619 HECT.
DOMAIN 35 38 POLY-GLY
SEQUENCE 619 AA; 70902 MW; 4CB2F8624A7B525 CRC64;
Query Match 61.7%; Score 2491.5; DB 1; Length 619;
Best Local Similarity 72.0%; Pred. No. 1.6e-170;
Matches 471; Conservative 59; Mismatches 75; Indels 49; Gaps 6;
109 DTGQRLDLCKLGNNDVVRGQIVVSLQSDRIGTGQGVVDCSLFNDLPDQWEERT 168
1 DTGQRLDLCKLNPSTDVAVRQIVVSLQTRIGGGSVVDCRGLLE----- 50
169 ASGRIOYLAHITTTQWRPTRPASEYSPGRFLSCFVDENTPISGTNGATCG----- 221
51 --GTAV-----EDSGRFLSCLMEEPAPYTDGTGAAGGNCRFV 89
222 --QSSDPLRAEVRVSQRHNYM-----SRTHLHTPPDLPEGYEORTTQGGVYPLHTQTG 275
90 ESPSQRLVLRQNPVEVQGLQTPQNRPHQSPPELPEGYEORTTQGGVYPLHTQTG 149
276 VSTWHPDPRVRLSNINCELGPLPPGWEIRNTATGRVYFVDHNNRTTQFTDPLRLSANLH 335
150 VSTWHPDPRVRLSNVNCDELGPLPPGWEIRNTATGRVYFVDHNNRTTQFTDPR----LH 205
336 LVLRNQKQKQKQ--QVVSCLPDDTECLTVPRYKRDIVQKILRQELSOQQPQAGHCR 394
206 HINMHCQLKEPSQPLQPLNEGSVEDELPQAYRERDLVQKLVRLHSLQQPQAGHCR 265
395 IEVSRREIPEESVROWKMRPKDLAKMLIKFERGEGLDYGVAREWMLYLLSHEMLNPPY 454
266 IEVSRREIPEESVROWKMRPKDLAKMLIKFERGEGLDYGVAREWMLYLLSHEMLNPPY 325

QY 455 GLFQVSRDDIYTLQINPDSSAVNPHELSYFHFVGRIMGVAPHGHYIDGGFTLPFYKOLLG 514
DB 326 GLFQVSTNDIYTLQINPDSSINPDHLSYFHFVGRIMGLAVFPHGHYINGGTVPFYKOLLG 385
QY 515 KSITIDDMELVDPDHLNSLVWILENDITGVLDHTFCVHNAYGELIOHELKPNKGKSPVN 574
DB 386 KPIQISDLESVDPELHKLWILENDITFVLDHTFCVHNAYGELIOHELKPNGRNVPVT 445
QY 575 ENKKEYVRLYVNWFLRGIEAQFALQKGFNEVIPHLLKTFDEKELELIICGLGKIDV 634
DB 446 ENKKEYVRLYVNWFLRGIEAQFALQKGFNEVIPHLLKTFDEKELELIICGLGKIDV 505
QY 635 NDWKYNTLKHCTPDNSNIVKFWKAVFFDEERARLLQFVTGSSRVPLOQFKALQGAAG 694
DB 506 NDWKSNTRLKHCVADSNIVRFWQAVETFDERRARLLQFVTGSTRVLPQFKALQGAAG 565
QY 695 PRLFTIHQIDACTNNLPKATCFNRIDIPPYSEYKLYEKLTAIETCGFAVE 748
DB 566 PRLFTIHLIDANTNLKPAHTCFNRIDIPPYSEYKLYEKLTAIETCGFAVE 619
RESULT 5
PUBL_SCHPO
ID PUBL_SCHPO STANDARD; PRT; 767 AA.
AC Q92452; O14454;
DT 01-NOV-1997 (Rel. 35; Created)
DT 15-DEC-1998 (Rel. 37; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Ubiquitin--protein ligase publi (EC 6.3.2.-).
GN PUBL OR SPAC11G7.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96205868; PubMed=8635463;
RA Nefsky B., Beach D.;
RT "Publ acts as an E6-AP-like protein ubiquitin ligase in the
RL degradation of cdc25."
RL EMBL J. 15:1301-1312(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=J227;
RX MEDLINE=97340937; PubMed=9197411;
RA Saleki R., Jia Z., Karagiannis J., Young P.G.;
RT "Tolerance of low pH in Schizosaccharomycetes pombe requires a
functioning publ ubiquitin ligase."
RL Mol. Gen. Genet. 254:520-528(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins M., Connor R., Cronin A., Davis P., Chillingworth T., Churcher C.M.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA McInerney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Welljars I., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljars I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Mosatl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

DR	InterPro; IPR002349; WW.
DR	InterPro; IPR001202; WW_Rep5_WMP.
DR	Pfam; PF00168; C2; 1.
DR	Pfam; PF00632; HECT; 1.
DR	Pfam; PF00397; WW; 3.
DR	PRINTS; PR00403; WWDOMAIN.
DR	SMART; SM00239; C2; 1.
DR	SMART; SM00119; HECTG; 1.
DR	SMART; SM00456; WW; 3.
DR	PROSITE; PS00499; C2_DOMAIN_1; 1.
DR	PROSITE; PS00004; C2_DOMAIN_2; 1.
DR	PROSITE; PS02037; HECT; 1.
DR	PROSITE; PS01159; WW_DOMAIN_1; 3.
DR	PROSITE; PS00020; WW_DOMAIN_2; 3.
KW	Ub1 conjugation pathway; Ligase; Repeat.
FT	DOMAIN 1 88
FT	DOMAIN 229
FT	DOMAIN 315
FT	DOMAIN 331
FT	DOMAIN 364
FT	DOMAIN 387
FT	DOMAIN 420
FT	DOMAIN 705
FT	DOMAIN 777
FT	BINDING 777
FT	VARIANT 733
FT	L--> S (IN RSP5-1; IMPAIRED IN UBIQUITIN-
FT	THIOESTER FORMATION AND CATALYSIS OF
FT	SUBSTRATE UBIQUITINATION).
FT	C->A: LOSS OF UBIQUITINATION.
FT	MUTAGEN 777
FT	SEQUENCE 809 AA; 91816 MW; 6F1836384479E70F CRC64;
Query Match	36.8%; Score 1484; DB 1; Length 809;
Best Local Similarity	37.1%; Pred.No. 2.6e-98;
Matches 324; Conservative 125; Mismatches 224; Indels 200; Gaps 20	
QY	11 PVKRLRTVLCAKNLVKKDFRLEDPFAKVVDVSGGCHSTDTVKNTLPKMNQHIDL-YI 69
DB	2 PSSISVKLVAESLYKEDVFERSDDPAVLITDGY-QTKSTSAAKTLNPNWNTEKFDDI 60
QY	70 GKSDSVTISVNNKHKKOGAGFLGCVRLLSNAINRKLD-----TGQR-----LDLC 118
DB	61 NENSILATIQVFDQKFF-KKKDQGFLGVNVVRGDVLGHLEDATSSGRPREETITRDLK 119
QY	119 KLGPDNDIVRGQIVW-----SLQSRDRIGTGGVVDCSR--- 153
DB	120 K-SNDGMVSGRIIVLVSKLPSSPHSQAPSGHTASSNTSSITRTNGHSTSTRHS 177
QY	154 -----LFDMND-----LPDGWEE 165
DB	178 TSHPSRGTAQAVESTLQSQTAAATNTATTSHRSTNSTSSATRQYSFSDQYGLPLPGWER 237
QY	166 RTASGRITQYLNLTRITQWERTPRASE-----YSSGRPLSCFVDENT 210
DB	238 RTDNFGRTYYVDHNTRITTWKRPDLQTEAERGQNQLNANTELRRHQHRTILPGSSDNS 297
QY	211 -----PISG-----TINGATCGOSSDPLRAERVRSQRHRNYMSTRHLHT 249
DB	298 SVTVQVGGSNI PVPNGAAAAAFATGGTTSLG-----LFDMND-----LPDGWEE 331
QY	250 PPDLPEGVEQRTQQGVYFLHTQGVSTVHDPRVPRDLNIN-----CBELGPL 299
DB	332 --ELPSEGWEQFTPEGRAYFVDHNTRITTTWDPFRQCYIRTYGTNTTIQQOPVSOLGPT 389
QY	300 PGWEIRNTARTGVYFVDHNRTTQTFDPLSANLHLVLRNQNLKDQQQQVVSCLCPDD 359
DB	390 PSGEWRLNTARVYFVDHNKTITWDDPLPSL-----DQ----- 426
QY	360 TECLIVPYKBDLVOKLIKILRQELSQQQPQ-----AGHCRIEVSRBEIFESTRWVMKP 415
DB	427 ----NVPQYKIDFRFKVIYFR-----SQPALRIPLQCCHKIKVRKNKIPEDAYQEIMRQFP 477
QY	416 XDLWKRLMKFEREGEDYGCVAREMLYLLSHEMLANPYGLFOYSRDDITYLTQINPDSAV 475
DB	478 EDLKKRLMIKIDBEGEDJOYGVGSREFFLLSHENFWNPFCLEFYSAIDNTYTQINPNSGI 537
QY	476 NPEHLSYFHFGVRIMGMAVFHGHYIDGGTLPFYKQLLGKSIITLDDMWELVDPDLHNSLV 535

538 NREHLNFKFICGVGLGVFHRFLDAFVGALYKMLKVKVLDMEGVDAEYVSLNW 597
 536 ILLENDITGVLDHTFCVENAVGEIIQHELKNGKSIPIVNEENKKEYVRLYNNRFLRGIE 595
 598 MLENSIDGVLDTTFADDERFGEVTVLKDGRNIEVTGCKEYVELYQWRIVDVQ 657
 596 APLALQKGFNEVIFQHLKTFDEKELELITCGLGKIDVNDKVNRLKHKTPDSNIVKW 655
 658 EQFKAFMGFNLIDEDLVTFDELELLGGTAEDIEDWKHTYRGYQSDDEVQW 717
 656 FWKAVEFFDEERARLLQVGTSSVPLQGFALQGAAGPLFTIHQIDACTNNLPKHAHT 715
 718 FWKCVSEWNEQARLLOFTTGTSTIPVNGFKDLQSGDPRFTIERAGE-VQQLPKSHT 776
 716 CNRIDIPYESYKLYEKLTAIBETCGFAVE 748
 777 CENRVDLPQVVDYDSMKQKLTIAVBETIGFQGE 809
 34 MOUSE
 NED4 MOUSE STANDARD; PRT; 957 AA.
 P46935;
 01-NOV-1995 (Rel. 32, Created)
 01-OCT-1996 (Rel. 34, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 NEDD-4 protein (EC 6.3.2.-) (Fragment).
 NEDD4 OR NEDD-4 OR NEDD4A.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Embryo;
 MEDLINE=92328780; PubMed=1378265;
 Kumar S, Tomooka Y., Noda M.;
 "Identification of a set of genes with developmentally down-regulated
 expression in the mouse brain.";
 Biochem. Biophys. Res. Commun. 185:1155-1161(1992).
 [2]
 REVISIONS.
 Kumar S.;
 Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
 AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
 THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES (BY
 SIMILARITY).
 -!- FUNCTION: INVOLVED IN THE EMBRYONIC DEVELOPMENT AND
 DIFFERENTIATION OF THE CENTRAL NERVOUS SYSTEM.
 -!- TISSUE SPECIFICITY: BRAIN.
 -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 thioester formation.
 -!- SIMILARITY: Contains 1 C2 domain.
 -!- SIMILARITY: Contains 3 WW domains.
 -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; D85414; BAA12803.1; -
 HSSP; Q13526; 1PIN.
 MGD; GI:97297; Nedd4.
 GO; GO:0005829; C:cytosol; IDA.
 GO; GO:000151; C:ubiquitin ligase complex; IPI.
 GO; GO:0005515; F:protein binding activity; IPI.

DR GO:0004842; F:ubiquitin-protein ligase activity; IDA.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR000569; HECT domain.
 DR InterPro; IPR001202; WW_Rsp5_WWP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00632; HECT; 1.
 DR Pfam; PF00397; WW; 3.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00119; HECTC; 1.
 DR SMART; SM00456; WW; 3.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS00004; C2_DOMAIN_2; 1.
 DR PROSITE; PS0237; HECT; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 3.
 DR PROSITE; PS00020; WW_DOMAIN_2; 3.
 KW Ubl conjugation pathway; Ligase; Repeat.
 FT NON_TER 1
 FT DOMAIN 150 236 C2 DOMAIN.
 FT DOMAIN 319 352 WW 1.
 FT DOMAIN 475 508 WW 2.
 FT DOMAIN 530 563 WW 3.
 FT DOMAIN 851 957 HECT.
 FT BINDING 924 924 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 957 AA; 109967 MW; 7955FIDF01879F28 CRC64;
 Query Match 33.6%; Score 1358; DB 1; Length 957;
 Best Local Similarity 37.0%; Pred. No. 3.4e-89;
 Matches 314; Conservative 125; Mismatches 249; Indels 160; Gaps 24;
 QY 14 LRLVLCALNKKVDFRLPDPFAKV---VDGSGQCHSTDTVNTLDPKNQHYDLY- 68
 DB 150 VRVVIAGIGLAKKDIIGASDPYRVTLVDPMGILTSVQTKIKSLNPKWEEI-LPR 208
 QY 69 -IGKSDSVTISVMNKKIKKQAGFLGCVR-----LASNAIN 105
 DB 209 VLPQRHRLFEVDPENLRTRD---FLGQVDVPLVPLTENPRMERPYTFKDFVLPFRSH 265
 QY 106 RLKGTGVRDLCKLGNPDNDTVRGQIVSLQSRDRIGTGGQVVD---CSRLL----FQN 157
 DB 266 KSRVGYRLKMTYLPKNGSED-----ENADQAELEPQVWVLDQFDAATHLPHPPEPS 319
 QY 158 DLPDQWEERTASGRIOYLNHITRTQWERPT-----RPAS 193
 DB 320 PLPQWEERQDVLGRTYVNHESRRQWKRPSPDDDLTDNDMDQLQARAFTRQIS 379
 QY 194 E-----YSGPRPLSCFVDENTPISG-----TNGATCGQ-- 222
 DB 380 EDVQDPNRSPEINWEIVREDENTYSGQAVQSPSGHIDVQTHLAEFEENTRLAVCGNPA 439
 QY 223 SSDRLAERRVRSQRHNYMSRTH-----LHTPDLPEGVQRITQOQGVVF 269
 DB 440 TSQP-----VTSSNHSRSGSLQTCIPEQPTLPVLLPTSSGLPPGWEEKQDDGRSY 493
 QY 270 LHTGTGVSTWHD-----RVPRDL-SNINCEELGPLPPGWENRNTATGRVYFVDHNN 320
 DB 494 VDHNSKTTWSKPTMQDDPRSKTFAHLRGKTDNDLGLPLPGWERTHTDGRVFFINHI 553
 QY 321 RTTQFTDPRLSANLHLVLFNQNLKQOQOQVSLCPDDTECLTVPRYKRLDVQKLILR 380
 DB 554 KKTQWEDPRL-----QNVAITGP-----AVP-YSRDYKRKFEPR 587
 QY 381 QELSQQOPQAGHCHRIEYSREIPEESYRQVMKRPKDLWK-RIMIKFRGEEGLDYGVAR 439
 DB 588 RKLKQTDIPNKPENKLRRAILEDVSRIMGVKRDLLKARLWIEFDGKGLDYGVAR 647
 QY 440 EWLVLHSEMLNPYGLFYQSRDDITYLQINPDASV-NPEHLSYFHFVGRIMGMAVFGH 498
 DB 648 EWPFLISKEMPNPYGLFEYSATDNTYLQINPNISGLCNEDHLSYFKFGRVAGMAYTHGK 707
 QY 499 YIDGGFTLPFYKQLLGSITLDDMELVDPLHNSLVILENDITGLDHTFCVEHNAYGE 558
 DB 708 LLDDGFFRPFYKMLQKLTILHDMESVDSEYSSSLKILENDPT-ELDLRFIDBELFQ 766

677 KLLDGFIRFFYKMLHKEITLHDMESVDSEYNSLRWILENDPT-ELDLRFLIIDEELPG 735
558 EIQHLEKNGKSIPIVNEENKEKYRLVNRFLRGIEAQFLALOKGFENEVIPOHLLKTF 617
736 OTHQHELKNGGSEIVVTKNKEKYTYLVQWRFNRIQKQMAAFKEGFFELIPQDLKIF 795
618 DEKELELIICGLGKIDVNDKVNTELKH-CTPDSNIVKFWKAVEFFDEERRARLLQFVT 676
796 DENELELLMCGLDVNDVNDREHTKYKGYNSANQVIOFWKAVLMDSEKIRLLQFVT 855
677 GSRVLPQFKALQGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPIYESYEKLYSKLL 736
856 GTSRVPNGFAELYGNGPSQSTVQW-GTPEKLPRAHTCFNRDLPPYESFEELMDKQ 914
737 TAJEETCGF 745
915 MALENTQGP 923

SULT 9
BI_RAT
P51593;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
DNA binding protein URE-B1 (EC 6.3.2.-).
UREB1.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain;
MEDLINE=95058008; PubMed=7968380;
Gu J., Ren K., Dubner R., Iadarola M.J.;
"Cloning of a DNA binding protein that is a tyrosine kinase substrate
and recognizes an upstream initiator-like sequence in the promoter of
the preprodynorphin gene.";
Brain Res. Mol. Brain Res. 24:77-88(1994).
-!- FUNCTION: MAY BE A DNA-BINDING TRANSCRIPTIONAL REGULATOR.
RECOGNIZES AN UPSTREAM INITIATOR-LIKE SEQUENCE IN THE PROMOTER OF
THE PREPRODYNORPHIN GENE.
-!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES (BY
SIMILARITY).
-!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
thioester formation.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U08214; AAA81950.1; -
InterPro; IPR000569; HECT_domain.
Pfam; PF00632; HECT; 1.
SMART; SM00119; HECTC; 1.
PROSITE; PS0237; HECT; 1.
Ub1 conjugation pathway; Ligase; Repeat; DNA-binding;
Transcription regulation.
DOMAIN 202 310 HECT.
BINDING 275 275 UBIQUITIN.
SEQUENCE 310 AA; 35640 MW; D6257F21F47C068E CRC64;
Query Match 15.5%; Score 626; DB 1; Length 310;

Best Local Similarity 44.1%; Pred. No. 1.5e-37;
Matches 135; Conservative
Indels 2; Gaps 2;
444 LLSHMLNPPYGLFOYSDDDIYLIQNPDSAVNPEHLSFHFVGRIMGMAVPHGHVIDGG 503
2 IISREMPMYALFRTSPGRVTRTSIHLPTVPTTTSATSSLAGRYVAKAVMTTALLECY 61
504 FTLFPYKOLLGKSIITLDDMELVDPDLHNSLVWILENDITGV-LDHTFCVHNAYGEIIQH 562
62 FTRFPYKHILGKSVRYTDMESDVHFYQGLVILENDVSTGLDYDLTFSVEQFGVCEVR 121
563 ELKNGKSIPIVNEENKEKYRLVNRFLRGIEAQFLALOKGFENEVIPOHLLKTFPEKEL 622
122 DLKPNGANILVTEENKKEYVHLCVQEMTGAIKQLAAPLEGFEYIIPKELISIFTEQEL 181
623 ELIICGLGKIDVNDKVNTELKHCTPDSNIVKFWKAVEFFDEERRARLLQFVTGSSRVP 682
182 ELLYTGLEFTIDIDLKNTYHKYQNSIQIOWFWRALRSFDQADRAKFLQFTGTGSKVP 241
683 LQGFALQGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPIYESYEKLYSKLLTAIET 742
242 LQGFALQGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPIYESYEKLYSKLLTAIET 742
743 C-GPAAV 747
302 SEGFGCL 307
RESULT 10
UE3A HUMAN
ID UE3A HUMAN STANDARD; PRT; 875 AA.
AC Q05086; P78355; Q93066;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3A (EC 6.3.2.-) (Oncogenic protein-
associated protein E6-AP) (Human papillomavirus E6-associated
protein).
DE UBE3A OR E6AP.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS I; II AND III).
RC TISSUE=Keratinocytes;
RX MEDLINE=9728825; PubMed=9143503;
RA Yamanoto Y., Huijbreghse J.M., Howley P.M.;
"The human E6-AP gene (UBE3A) encodes three potential protein isoforms
generated by differential splicing.";
RL Genomics 41:263-266(1997).
RN [2]
RP SEQUENCE OF 11-875 FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Keratinocytes;
RX MEDLINE=93140775; PubMed=8380895;
RA Huijbreghse J.M., Scheffner M., Howley P.M.;
"Cloning and expression of the cDNA for E6-AP, a protein that
mediates the interaction of the human papillomavirus E6 oncoprotein
with p53.";
RL Mol. Cell. Biol. 13:775-784(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM I), AND VARIANT VAL-290.
RC TISSUE=Fetal brain;
RX MEDLINE=97141921; PubMed=8988171;
RA Kishino T., Lalonde M., Wagstaff J.;
"UBE3A/E6-AP mutations cause Angelman syndrome.";
RL Nat. Genet. 15:70-73(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM I), AND VARIANTS TYR-44; THR-201 AND
VAL-290.
RX MEDLINE=97141922; PubMed=8988172;
RA Matsuura T., Sutcliffe J.S., Fang P., Galjaard R.-J., Jiang Y.-H.,
Benton C.S., Rommens J.M., Beaudet A.L.;

"De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome.";
Nat. Genet. 15:74-77(1997).
[15]
VARIANT AS ILE-826 INS, AND VARIANTS HIS-62; THR-201 AND PRO-372.
MEDLINE=98254455; PubMed=9885605;
Maizac P., Webber H., Moncla A., Graham J.M. Jr., Kukolich M.,
Williams C., Pagon R.A., Ramesh L.A., Kishino T., Wagstaff J.,
"Mutation analysis of UBE3A in Angelman syndrome patients.";
Am. J. Hum. Genet. 62:1353-1360(1998).
[16]
CHARACTERIZATION.
MEDLINE=98351563; PubMed=9688277;
Nuber U., Schwarz S.E., Scheffner M.;
"The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as
its own substrate";
Eur. J. Biochem. 254:643-649(1998).
-!- FUNCTION: INTERACTS WITH THE E6 PROTEIN OF THE CANCER-ASSOCIATED
HUMAN PAPILLOMAVIRUS TYPES 16 AND 18. THE E6/E6-AP COMPLEX BINDS
TO AND TARGETS THE P53 TUMOR-SUPPRESSOR PROTEIN FOR UBIQUITIN-
MEDIATED PROTEOLYSIS. IT IS AN E3 UBIQUITIN-PROTEIN LIGASE WHICH
ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE
FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO
TARGETED SUBSTRATES. IT CAN TARGET ITSELF FOR UBIQUITINATION IN
VITRO AND EFFICIENTLY PROMOTES ITS OWN DEGRADATION IN VIVO. IT
APPEARS THAT ONLY UNMODIFIED E6-AP MOLECULES CAN BIND EFFICIENTLY
TO P53 IN THE PRESENCE OF THE HPV E6 ONCOPROTEIN.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Name=I1;
IsoId=Q05086-1; Sequence=Displayed;
Name=I;
IsoId=Q05086-2; Sequence=VSP_006705;
Name=III;
IsoId=Q05086-3; Sequence=VSP_006706;
-!- DISEASE: DEFECTS IN UBE3A ARE A CAUSE OF ANGELMAN SYNDROME (AS),
ALSO KNOWN AS 'HAPPY PUPPET SYNDROME'. IT IS CHARACTERIZED BY
FEATURES OF SEVERE MOTOR AND INTELLECTUAL RETARDATION,
MICROCEPHALY, ATAXIA, FREQUENT JERKY LIMB MOVEMENTS AND FLAPPING
OF THE ARMS AND HANDS, HYPTONIA, HYPERACTIVITY, HYPOSEGMENTATION,
SEIZURES, ABSENCE OF SPEECH, FREQUENT SMILING AND EPISODES OF
PAROXYSMAL LAUGHTER, AND AN UNUSUAL FACIES CHARACTERIZED BY
MACROSTOMIA, A LARGE MANDIBLE AND OPEN-MOUTHED EXPRESSION, A GREAT
PROPENSITY FOR PROTRUDING THE TONGUE ('TONGUE THRUSTING'), AND AN
OCCUPATIONAL GROOVE.
-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
THIOESTER FORMATION.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-!- SIMILARITY: SOME, IN THE C-TERMINAL HALF, TO RAT 100 kDa PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/announcement/>
or send an email to license@ebi-sib.ch).

EMBL; X98021; CAA66653.1; JOINED.
EMBL; X98027; CAA66653.1; JOINED.
EMBL; X98022; CAA66653.1; JOINED.
EMBL; X98023; CAA66653.1; JOINED.
EMBL; X98024; CAA66653.1; JOINED.
EMBL; X98025; CAA66653.1; JOINED.
EMBL; X98026; CAA66653.1; JOINED.
EMBL; X98028; CAA66653.1; JOINED.
EMBL; X98029; CAA66653.1; JOINED.
EMBL; X98030; CAA66653.1; JOINED.
EMBL; X98033; CAA66656.1; -.
EMBL; X98031; CAA66654.1; -.
EMBL; X98032; CAA66655.1; -.

DR EMBL; L07557; AAA35542.1; ALT_INIT.
DR EMBL; AF016708; AAB69154.1; -.
DR EMBL; AF016703; AAB69154.1; JOINED.
DR EMBL; AF016704; AAB69154.1; JOINED.
DR EMBL; AF016705; AAB69154.1; JOINED.
DR EMBL; AF016706; AAB69154.1; JOINED.
DR EMBL; AF016707; AAB69154.1; JOINED.
DR EMBL; U84404; AAB49301.1; -.
DR PDB; 1C42; 17-NOV-99.
DR PDB; 1DSF; 17-NOV-99.
DR PDB; 1EQX; 28-FEB-01.
DR Gene; HGNC:12496; UBE3A.
DR MIN; 601623; -.
DR MIN; 105830; -.
DR GO; GO:0004840; Fubiquitin conjugating enzyme activity; TAS.
DR GO; GO:0007420; P:brain development; TAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS0237; HECT; 1.
KW Nuclear protein; ubl conjugation pathway; Ligase;
KW Alternative splicing; Disease mutation; Polymorphism; 3D-structure.
FT DOMAIN 394 399
FT BINDING 776 875
FT BINDING 843 843
FT BINDING 1 23
FT VARSPLIC 1 10
FT VARSPLIC 44 44
FT VARIANT 62 62
FT VARIANT 201 201
FT VARIANT 290 290
FT VARIANT 372 372
FT VARIANT 826 826
FT CONFLICT 359 359
FT CONFLICT 423 423
FT CONFLICT 647 649
FT CONFLICT 669 669
FT CONFLICT 686 686
SQ SEQUENCE 875 AA; 100645 MW; 667925F1956670AA CRC64;
Query Match 13.3%; Score 539; DB 1; Length 875;
Best Local Similarity 25.1%; Pred. No. 1.1e-20;
Matches 183; Conservative 117; Mismatches 277; Indels 152; Gaps 22;
QY 105 NLRKDTGYQLDLCKLGNPNDVTRGQIVVLSQSDRIGTG-----GQVVDGS 152
D 212 SRIGDSQGNLNQKLGPDVSDIDAIR-RVYTRLLSNEKIEAFNALVYLSNPVEDC 270
QY 153 RLFPNDLPDNGEERTASGIVLNHITTTQWERTPRASEYSSSPGRLSCFVDENTPI 212
D 271 LTYEN-----VYSRDPNYLNF--IIGMENNLHSPEYLEMALPLFCRAMSKLPL 318
QY 213 SGT-----NGATCGQSSDPLAERR----- 232
D 319 AAOGLRLMSKYNADQIRRMWETFQOLITYKISNEFNSRLNVNDDDAIWAASKLKMV 378
QY 233 -----VRSQRHNTMSRTHLHTPPDLP-----GYEORTQOQGVYFHTQTGVSTWH 280
D 379 YYANVVGGEVDTHNNEEDDEEPPESELTLQELLGEEERNKKGPRVDFLTGELGVKT-- 436
QY 281 DPRVPRDLNSINCEELGFLPPGWEIRNTATGRV-----YFVDHNNRTTQFIDPRL- 330
D 437 -----LDCKK--PLIPFEEFNEPLNEVLEMDKDYTFKVTENKFSFMTCTFFL 484

523 SLVQGOQLNPFYTLKVRDHIIDDLALVRLAMENPADLKKQLYVEBEGQVDEGVGS 582
439 REMWYLLSHMLNPNYGLFOYSRDDIYTL-QINPDSAVNPEHLSHFVFRIMGMAVPHG 497
583 KEFFQLVVEIFNPDIQMETY-DEATKLFWNPSSF---ETEGQFTLIG-ILGLAIYNN 636
498 HYIDGGFTLPFYKOLKGSITLDDMBELVDPDLHNSLVWILEND-----ITGVLDHTF 549
637 CILDVHPMVVYKLMKGGTFDDLGSDHPVLQSLKDLLEYEGSVEDDMWITFISQT- 695
550 CVBHNAYGEIIOHELKPNKSIIPVNEENKVEYVLYVNRFLRGIERAQLAQKGFNEVI 609
696 ---DLFGNPMYDLKENGKIBITNENKEFEVISDYILNKSVKQPKAFRRGFHMVT 751
610 POHLK-TPEKELELIICGLKIDVNDKVNRLK-HCTPDNSIVKMFKAVERFEDEER 667
752 NESPLKLFPEIEIILICGRNLDFOALETTEYDGGTVRESVIRFELVHSFTDEQ 811
668 RARLQFVTSRVPLOG---FKALQGAAGPRLFTTHQIDACTNNLPKHAHTCFNRIDIPP 724
812 KRLFLLETTGTDRAPVGLGKLKMITAKNGPD-----TERLPTSHTCFNVLLLPE 861
725 YESYEKLYEKLTAIBETCGFAV 747
862 YSSKEKLVNRLKAITYAKFGM 884

RESULT 12

HL4 YEAST
P40985; HUL4 YEAST STANDARD; PRT; 892 AA.
01-FEB-1995 (Rel. 31, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Probable ubiquitin-protein ligase HUL4 (EC 6.3.2.-):
HUL4 OR YR036C OR J1608.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
STRAIN=S288C;
MEDLINE=95397595; PubMed=7668047;
Huang M.-E., Chuat J.-C., Galibert F.;
"Analysis of a 42.5 kb DNA sequence of chromosome X reveals three
tRNA genes and 14 new open reading frames including a gene most
probably belonging to the family of ubiquitin-protein ligases";
Yeast 11:775-781(1995).
[2]
SEQUENCE OF 362-892 FROM N.A.
MEDLINE=95043431; PubMed=7957102;
van Gool A.J., Verhage R., Swagemakers S.M.A., van de Putte P.,
Brouwer J., Troelstra C., Bootsma D., Hoeijmakers J.H.J.;
"RAD26, the functional S. cerevisiae homolog of the Cockayne syndrome
B gene ERCC6";
EMBO J. 13:5361-5369(1994).
[3]
GENE NAME, AND GENE DISRUPTION.
MEDLINE=99077972; PubMed=9858558;
Wang G., Yang J., Huibregtse J.M.;
"Functional domains of the repress ubiquitin-protein ligase";
Mol. Cell. Biol. 19:342-352(1999).
-!- FUNCTION: PROBABLE E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS
UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A
THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED
SUBSTRATES. NON ESSENTIAL.
-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
THIOESTER FORMATION (BY SIMILARITY).
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; L36344; AAA88738.1; -
DR EMBL; Z49536; CA889563.1; -
DR EMBL; X81635; CA857291.1; -
DR PIR; S57055; S57055.
DR SGD; S0003797; YJR036C.
DR InterPro; IPR000559; HECT_domain.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
KM Ub1 conjugation pathway; Ligase.
FT DOMAIN 792 892
FT BINDING 860 860
FT CONFLICT 362 362
FT CONFLICT 410 410
FT CONFLICT 418 418
FT CONFLICT 514 519
FT CONFLICT 589 607
FT CONFLICT 705 705
FT CONFLICT 723 724
FT CONFLICT 723 724
SQ SEQUENCE 892 AA; 103456 MW; 35EF9A8DDA92BD84 CRC64;
Query Match 11.4%; Score 460.5; DB 1; Length 892;
Best Local Similarity 24.5%; Pred. No. 4,7e-25;
Matches 172; Conservative 114; Mismatches 266; Indels 151; Gaps 25;
Qy 96 CVLLSNAINRLKDTGYQLDLCKLPGNDNDTVRQIV---VSLQSRDRIGTGQVDCS 152
Db 285 CIGYLSNLSTK--TSQQLIQSLRRTPTDNTFSYQVEILNLYINQ-----FS 328
Qy 153 RLFDNDLPQWEERRATASRIQYLNHITRTTCWERPTPASEYSPGRLPSCFVDENTPI 212
Db 329 RLLSNEL-----SNRTAKNNVPEDEM--RSRLRHHTTGHFUST-RPIS----- 371
Qy 213 SGTNGATCGQSDPRLA--ERRVRSQ---RHRNYSMTSLHTPTDLPREGYQRTTQGOV 267
Db 372 -----AQSDNKGSGTFHPVNNKMKFPQYEDNHHSAAKLTFIY----- 413
Qy 268 YFLHTGTGVSTWHDPRVPRDLNINCELGPLPPGWEIRNATGRVYFVDHNNRTQTFD 327
Db 414 YVANT-----RRNGRGALSISQSFYNTLDFID 440
Qy 328 PRLSANLH---LVLNRONLKDOQQQVWSLCPDDTECLTVPRYKRDVLQKLI----- 378
Db 441 YKQDFD-HWRGVAQKTKMQLIEWGNSTTK-----KCFSECKYFFLLSLGKISIMEY 493
Qy 379 -LRQELSQOQQA-----GHCRIEVSREIPEESYRQVQMKRPDLWRLMIK 425
Db 494 EIRIMEHEAEQAFILSLDKGKSDVDFYFKIVKVRDVISHDLSR-CIKEHQGDLLKSLRIE 552
Qy 426 FRCEGLDYGVAWEWLYLLSHEMLNPFYGLFQYSRDDIYT-LQINP---DSAVNPEHL 480
Db 553 FVNEPGIDAGLGRKEWFFLLTKSLFPNPNGLFIYIKESRSWFALDPNPFKSGKNSQL 612
Qy 481 SYPHFVGRINGMAVPHGYIDGGFTLPFYKOLKGSITLDDMBELVDPDLHNSLVWIL--- 537
Db 613 ELVYLFQVWGLAIFNSTILDLPFKALYKLCSEPLSFEDYSELFETSRNLKMLNYT 672
Qy 538 ENDITGVLDHTF-----CVEHNAIGEIIQHELKPNKSIIPVNEENKVEYVLYV 587
Db 673 EDNPFDFVSLTFETTVRNANNWILNDSKSKKEYVTVELCENGRNVPITQSNKHEFVKNWE 732
Qy 588 WRFLRGIEAQLALQGFNEVIPQ-HLAKTFDEKELELIICGLKIDVNDMK-----VNTR 642
Db 733 FYLEKSIETQYNKVFSGFKRFAECNSIKLFNSELERLVCGDEBQTKDFPKSLRSVTKY 792

```
643 LKCTPDNSIVKWFKAVERFFDEERARILQVGTSSRVELQGFALQGAAGPRLTIHQ 702
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
793 VGFSDSDRAVCWFELISWDFYQKLLQVFTASDRIPATGISTIFPKIS--LLGSHD 850
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
703 IDACTNPLPKAHTCFNRIDIPPPYSEYKELYKLLTAIBETQGF 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
851 SD----DLPLAHTCFNEICLWNYSSKKLELKLWAINESGY 889
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

SULT 13
R3_HUMAN STANDARD; PRT; 1050 AA.
Q15034;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
HECT domain and RCC1-like domain protein 3.
HERC3 OR KIAA0032.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Bone marrow;
MEDLINE=96051387; PubMed=7584026;
Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y.,
Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
"Prediction of the coding sequences of unidentified human genes. I.
The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
analysis of randomly sampled cDNA clones from human immature myeloid
cell line KG-1";
DNA Res. 1:27-35 (1994).
[2]
CHARACTERIZATION.
MEDLINE=21099818; PubMed=11163799;
Cruz C., Ventura F., Bartrons R., Rosa J.L.;
"HERC3 binding to and regulation by ubiquitin.";
FEBS Lett. 488:74-80(2001).
-1- SUBCELLULAR LOCATION: Cytoplasmic. Also found in vesicular-like
structures.
-1- PTM: Substrate of ubiquitination and is degraded by the
proteasome.
-1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-1- SIMILARITY: Contains 7 RCC1 repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D25215; BAA04945.1; -.
Genew; HGNC:4876; HERC3.
MIM; 605200; -.
InterPro; IPR000569; HECT domain.
InterPro; IPR000408; Reg_chr_condens.
Pfam; PF00632; HECT; 1.
Pfam; PF00415; RCC1; 4.
PRINTS; PR00633; RCCNDSATION.
SMART; SM00119; HECTC; 1.
PROSITE; PS02037; HECT; 1.
PROSITE; PS00625; RCC1_1; FALSE_NEG.
PROSITE; PS00625; RCC1_2; 4.
PROSITE; PS00625; RCC1_3; 7.
PROSITE; PS00625; RCC1_4; 7.
Ub1 conjugation pathway; Ub1 conjugation; Repeat.
REPEAT 1 51
REPEAT 2 51
REPEAT 3 101
REPEAT 4 101
REPEAT 5 101
REPEAT 6 101
REPEAT 7 101
REPEAT 8 101
REPEAT 9 101
REPEAT 10 101
REPEAT 11 101
REPEAT 12 101
REPEAT 13 101
REPEAT 14 101
REPEAT 15 101
REPEAT 16 101
REPEAT 17 101
REPEAT 18 101
REPEAT 19 101
REPEAT 20 101
REPEAT 21 101
REPEAT 22 101
REPEAT 23 101
REPEAT 24 101
REPEAT 25 101
REPEAT 26 101
REPEAT 27 101
REPEAT 28 101
REPEAT 29 101
REPEAT 30 101
REPEAT 31 101
REPEAT 32 101
REPEAT 33 101
REPEAT 34 101
REPEAT 35 101
REPEAT 36 101
REPEAT 37 101
REPEAT 38 101
REPEAT 39 101
REPEAT 40 101
REPEAT 41 101
REPEAT 42 101
REPEAT 43 101
REPEAT 44 101
REPEAT 45 101
REPEAT 46 101
REPEAT 47 101
REPEAT 48 101
REPEAT 49 101
REPEAT 50 101
REPEAT 51 101
REPEAT 52 101
REPEAT 53 101
REPEAT 54 101
REPEAT 55 101
REPEAT 56 101
REPEAT 57 101
REPEAT 58 101
REPEAT 59 101
REPEAT 60 101
REPEAT 61 101
REPEAT 62 101
REPEAT 63 101
REPEAT 64 101
REPEAT 65 101
REPEAT 66 101
REPEAT 67 101
REPEAT 68 101
REPEAT 69 101
REPEAT 70 101
REPEAT 71 101
REPEAT 72 101
REPEAT 73 101
REPEAT 74 101
REPEAT 75 101
REPEAT 76 101
REPEAT 77 101
REPEAT 78 101
REPEAT 79 101
REPEAT 80 101
REPEAT 81 101
REPEAT 82 101
REPEAT 83 101
REPEAT 84 101
REPEAT 85 101
REPEAT 86 101
REPEAT 87 101
REPEAT 88 101
REPEAT 89 101
REPEAT 90 101
REPEAT 91 101
REPEAT 92 101
REPEAT 93 101
REPEAT 94 101
REPEAT 95 101
REPEAT 96 101
REPEAT 97 101
REPEAT 98 101
REPEAT 99 101
REPEAT 100 101
REPEAT 101 101
REPEAT 102 101
REPEAT 103 101
REPEAT 104 101
REPEAT 105 101
REPEAT 106 101
REPEAT 107 101
REPEAT 108 101
REPEAT 109 101
REPEAT 110 101
REPEAT 111 101
REPEAT 112 101
REPEAT 113 101
REPEAT 114 101
REPEAT 115 101
REPEAT 116 101
REPEAT 117 101
REPEAT 118 101
REPEAT 119 101
REPEAT 120 101
REPEAT 121 101
REPEAT 122 101
REPEAT 123 101
REPEAT 124 101
REPEAT 125 101
REPEAT 126 101
REPEAT 127 101
REPEAT 128 101
REPEAT 129 101
REPEAT 130 101
REPEAT 131 101
REPEAT 132 101
REPEAT 133 101
REPEAT 134 101
REPEAT 135 101
REPEAT 136 101
REPEAT 137 101
REPEAT 138 101
REPEAT 139 101
REPEAT 140 101
REPEAT 141 101
REPEAT 142 101
REPEAT 143 101
REPEAT 144 101
REPEAT 145 101
REPEAT 146 101
REPEAT 147 101
REPEAT 148 101
REPEAT 149 101
REPEAT 150 101
REPEAT 151 101
REPEAT 152 101
REPEAT 153 101
REPEAT 154 101
REPEAT 155 101
REPEAT 156 101
REPEAT 157 101
REPEAT 158 101
REPEAT 159 101
REPEAT 160 101
REPEAT 161 101
REPEAT 162 101
REPEAT 163 101
REPEAT 164 101
REPEAT 165 101
REPEAT 166 101
REPEAT 167 101
REPEAT 168 101
REPEAT 169 101
REPEAT 170 101
REPEAT 171 101
REPEAT 172 101
REPEAT 173 101
REPEAT 174 101
REPEAT 175 101
REPEAT 176 101
REPEAT 177 101
REPEAT 178 101
REPEAT 179 101
REPEAT 180 101
REPEAT 181 101
REPEAT 182 101
REPEAT 183 101
REPEAT 184 101
REPEAT 185 101
REPEAT 186 101
REPEAT 187 101
REPEAT 188 101
REPEAT 189 101
REPEAT 190 101
REPEAT 191 101
REPEAT 192 101
REPEAT 193 101
REPEAT 194 101
REPEAT 195 101
REPEAT 196 101
REPEAT 197 101
REPEAT 198 101
REPEAT 199 101
REPEAT 200 101
REPEAT 201 101
REPEAT 202 101
REPEAT 203 101
REPEAT 204 101
REPEAT 205 101
REPEAT 206 101
REPEAT 207 101
REPEAT 208 101
REPEAT 209 101
REPEAT 210 101
REPEAT 211 101
REPEAT 212 101
REPEAT 213 101
REPEAT 214 101
REPEAT 215 101
REPEAT 216 101
REPEAT 217 101
REPEAT 218 101
REPEAT 219 101
REPEAT 220 101
REPEAT 221 101
REPEAT 222 101
REPEAT 223 101
REPEAT 224 101
REPEAT 225 101
REPEAT 226 101
REPEAT 227 101
REPEAT 228 101
REPEAT 229 101
REPEAT 230 101
REPEAT 231 101
REPEAT 232 101
REPEAT 233 101
REPEAT 234 101
REPEAT 235 101
REPEAT 236 101
REPEAT 237 101
REPEAT 238 101
REPEAT 239 101
REPEAT 240 101
REPEAT 241 101
REPEAT 242 101
REPEAT 243 101
REPEAT 244 101
REPEAT 245 101
REPEAT 246 101
REPEAT 247 101
REPEAT 248 101
REPEAT 249 101
REPEAT 250 101
REPEAT 251 101
REPEAT 252 101
REPEAT 253 101
REPEAT 254 101
REPEAT 255 101
REPEAT 256 101
REPEAT 257 101
REPEAT 258 101
REPEAT 259 101
REPEAT 260 101
REPEAT 261 101
REPEAT 262 101
REPEAT 263 101
REPEAT 264 101
REPEAT 265 101
REPEAT 266 101
REPEAT 267 101
REPEAT 268 101
REPEAT 269 101
REPEAT 270 101
REPEAT 271 101
REPEAT 272 101
REPEAT 273 101
REPEAT 274 101
REPEAT 275 101
REPEAT 276 101
REPEAT 277 101
REPEAT 278 101
REPEAT 279 101
REPEAT 280 101
REPEAT 281 101
REPEAT 282 101
REPEAT 283 101
REPEAT 284 101
REPEAT 285 101
REPEAT 286 101
REPEAT 287 101
REPEAT 288 101
REPEAT 289 101
REPEAT 290 101
REPEAT 291 101
REPEAT 292 101
REPEAT 293 101
REPEAT 294 101
REPEAT 295 101
REPEAT 296 101
REPEAT 297 101
REPEAT 298 101
REPEAT 299 101
REPEAT 300 101
REPEAT 301 101
REPEAT 302 101
REPEAT 303 101
REPEAT 304 101
REPEAT 305 101
REPEAT 306 101
REPEAT 307 101
REPEAT 308 101
REPEAT 309 101
REPEAT 310 101
REPEAT 311 101
REPEAT 312 101
REPEAT 313 101
REPEAT 314 101
REPEAT 315 101
REPEAT 316 101
REPEAT 317 101
REPEAT 318 101
REPEAT 319 101
REPEAT 320 101
REPEAT 321 101
REPEAT 322 101
REPEAT 323 101
REPEAT 324 101
REPEAT 325 101
REPEAT 326 101
REPEAT 327 101
REPEAT 328 101
REPEAT 329 101
REPEAT 330 101
REPEAT 331 101
REPEAT 332 101
REPEAT 333 101
REPEAT 334 101
REPEAT 335 101
REPEAT 336 101
REPEAT 337 101
REPEAT 338 101
REPEAT 339 101
REPEAT 340 101
REPEAT 341 101
REPEAT 342 101
REPEAT 343 101
REPEAT 344 101
REPEAT 345 101
REPEAT 346 101
REPEAT 347 101
REPEAT 348 101
REPEAT 349 101
REPEAT 350 101
REPEAT 351 101
REPEAT 352 101
REPEAT 353 101
REPEAT 354 101
REPEAT 355 101
REPEAT 356 101
REPEAT 357 101
REPEAT 358 101
REPEAT 359 101
REPEAT 360 101
REPEAT 361 101
REPEAT 362 101
REPEAT 363 101
REPEAT 364 101
REPEAT 365 101
REPEAT 366 101
REPEAT 367 101
REPEAT 368 101
REPEAT 369 101
REPEAT 370 101
REPEAT 371 101
REPEAT 372 101
REPEAT 373 101
REPEAT 374 101
REPEAT 375 101
REPEAT 376 101
REPEAT 377 101
REPEAT 378 101
REPEAT 379 101
REPEAT 380 101
REPEAT 381 101
REPEAT 382 101
REPEAT 383 101
REPEAT 384 101
REPEAT 385 101
REPEAT 386 101
REPEAT 387 101
REPEAT 388 101
REPEAT 389 101
REPEAT 390 101
REPEAT 391 101
REPEAT 392 101
REPEAT 393 101
REPEAT 394 101
REPEAT 395 101
REPEAT 396 101
REPEAT 397 101
REPEAT 398 101
REPEAT 399 101
REPEAT 400 101
REPEAT 401 101
REPEAT 402 101
REPEAT 403 101
REPEAT 404 101
REPEAT 405 101
REPEAT 406 101
REPEAT 407 101
REPEAT 408 101
REPEAT 409 101
REPEAT 410 101
REPEAT 411 101
REPEAT 412 101
REPEAT 413 101
REPEAT 414 101
REPEAT 415 101
REPEAT 416 101
REPEAT 417 101
REPEAT 418 101
REPEAT 419 101
REPEAT 420 101
REPEAT 421 101
REPEAT 422 101
REPEAT 423 101
REPEAT 424 101
REPEAT 425 101
REPEAT 426 101
REPEAT 427 101
REPEAT 428 101
REPEAT 429 101
REPEAT 430 101
REPEAT 431 101
REPEAT 432 101
REPEAT 433 101
REPEAT 434 101
REPEAT 435 101
REPEAT 436 101
REPEAT 437 101
REPEAT 438 101
REPEAT 439 101
REPEAT 440 101
REPEAT 441 101
REPEAT 442 101
REPEAT 443 101
REPEAT 444 101
REPEAT 445 101
REPEAT 446 101
REPEAT 447 101
REPEAT 448 101
REPEAT 449 101
REPEAT 450 101
REPEAT 451 101
REPEAT 452 101
REPEAT 453 101
REPEAT 454 101
REPEAT 455 101
REPEAT 456 101
REPEAT 457 101
REPEAT 458 101
REPEAT 459 101
REPEAT 460 101
REPEAT 461 101
REPEAT 462 101
REPEAT 463 101
REPEAT 464 101
REPEAT 465 101
REPEAT 466 101
REPEAT 467 101
REPEAT 468 101
REPEAT 469 101
REPEAT 470 101
REPEAT 471 101
REPEAT 472 101
REPEAT 473 101
REPEAT 474 101
REPEAT 475 101
REPEAT 476 101
REPEAT 477 101
REPEAT 478 101
REPEAT 479 101
REPEAT 480 101
REPEAT 481 101
REPEAT 482 101
REPEAT 483 101
REPEAT 484 101
REPEAT 485 101
REPEAT 486 101
REPEAT 487 101
REPEAT 488 101
REPEAT 489 101
REPEAT 490 101
REPEAT 491 101
REPEAT 492 101
REPEAT 493 101
REPEAT 494 101
REPEAT 495 101
REPEAT 496 101
REPEAT 497 101
REPEAT 498 101
REPEAT 499 101
REPEAT 500 101
REPEAT 501 101
REPEAT 502 101
REPEAT 503 101
REPEAT 504 101
REPEAT 505 101
REPEAT 506 101
REPEAT 507 101
REPEAT 508 101
REPEAT 509 101
REPEAT 510 101
REPEAT 511 101
REPEAT 512 101
REPEAT 513 101
REPEAT 514 101
REPEAT 515 101
REPEAT 516 101
REPEAT 517 101
REPEAT 518 101
REPEAT 519 101
REPEAT 520 101
REPEAT 521 101
REPEAT 522 101
REPEAT 523 101
REPEAT 524 101
REPEAT 525 101
REPEAT 526 101
REPEAT 527 101
REPEAT 528 101
REPEAT 529 101
REPEAT 530 101
REPEAT 531 101
REPEAT 532 101
REPEAT 533 101
REPEAT 534 101
REPEAT 535 101
REPEAT 536 101
REPEAT 537 101
REPEAT 538 101
REPEAT 539 101
REPEAT 540 101
REPEAT 541 101
REPEAT 542 101
REPEAT 543 101
REPEAT 544 101
REPEAT 545 101
REPEAT 546 101
REPEAT 547 101
REPEAT 548 101
REPEAT 549 101
REPEAT 550 101
REPEAT 551 101
REPEAT 552 101
REPEAT 553 101
REPEAT 554 101
REPEAT 555 101
REPEAT 556 101
REPEAT 557 101
REPEAT 558 101
REPEAT 559 101
REPEAT 560 101
REPEAT 561 101
REPEAT 562 101
REPEAT 563 101
REPEAT 564 101
REPEAT 565 101
REPEAT 566 101
REPEAT 567 101
REPEAT 568 101
REPEAT 569 101
REPEAT 570 101
REPEAT 571 101
REPEAT 572 101
REPEAT 573 101
REPEAT 574 101
REPEAT 575 101
REPEAT 576 101
REPEAT 577 101
REPEAT 578 101
REPEAT 579 101
REPEAT 580 101
REPEAT 581 101
REPEAT 582 101
REPEAT 583 101
REPEAT 584 101
REPEAT 585 101
REPEAT 586 101
REPEAT 587 101
REPEAT 588 101
REPEAT 589 101
REPEAT 590 101
REPEAT 591 101
REPEAT 592 101
REPEAT 593 101
REPEAT 594 101
REPEAT 595 101
REPEAT 596 101
REPEAT 597 101
REPEAT 598 101
REPEAT 599 101
REPEAT 600 101
REPEAT 601 101
REPEAT 602 101
REPEAT 603 101
REPEAT 604 101
REPEAT 605 101
REPEAT 606 101
REPEAT 607 101
REPEAT 608 101
REPEAT 609 101
REPEAT 610 101
REPEAT 611 101
REPEAT 612 101
REPEAT 613 101
REPEAT 614 101
REPEAT 615 101
REPEAT 616 101
REPEAT 617 101
REPEAT 618 101
REPEAT 619 101
REPEAT 620 101
REPEAT 621 101
REPEAT 622 101
REPEAT 623 101
REPEAT 624 101
REPEAT 625 101
REPEAT 626 101
REPEAT 627 101
REPEAT 628 101
REPEAT 629 101
REPEAT 630 101
REPEAT 631 101
REPEAT 632 101
REPEAT 633 101
REPEAT 634 101
REPEAT 635 101
REPEAT 636 101
REPEAT 637 101
REPEAT 638 101
REPEAT 639 101
REPEAT 640 101
REPEAT 641 101
REPEAT 642 101
REPEAT 643 101
REPEAT 644 101
REPEAT 645 101
REPEAT 646 101
REPEAT 647 101
REPEAT 648 101
REPEAT 649 101
REPEAT 650 101
REPEAT 651 101
REPEAT 652 101
REPEAT 653 101
REPEAT 654 101
REPEAT 655 101
REPEAT 656 101
REPEAT 657 101
REPEAT 658 101
REPEAT 659 101
REPEAT 660 101
REPEAT 661 101
REPEAT 662 101
REPEAT 663 101
REPEAT 664 101
REPEAT 665 101
REPEAT 666 101
REPEAT 667 101
REPEAT 668 101
REPEAT 669 101
REPEAT 670 101
REPEAT 671 101
REPEAT 672 101
REPEAT 673 101
REPEAT 674 101
REPEAT 675 101
REPEAT 676 101
REPEAT 677 101
REPEAT 678 101
REPEAT 679 101
REPEAT 680 101
REPEAT 681 101
REPEAT 682 101
REPEAT 683 101
REPEAT 684 101
REPEAT 685 101
REPEAT 686 101
REPEAT 687 101
REPEAT 688 101
REPEAT 689 101
REPEAT 690 101
REPEAT 691 101
REPEAT 692 101
REPEAT 693 101
REPEAT 694 101
REPEAT 695 101
REPEAT 696 101
REPEAT 697 101
REPEAT 698 101
REPEAT 699 101
REPEAT 700 101
REPEAT 701 101
REPEAT 702 101
REPEAT 703 101
REPEAT 704 101
REPEAT 705 101
REPEAT 706 101
REPEAT 707 101
REPEAT 708 101
REPEAT 709 101
REPEAT 710 101
REPEAT 711 101
REPEAT 712 101
REPEAT 713 101
REPEAT 714 101
REPEAT 715 101
REPEAT 716 101
REPEAT 717 101
REPEAT 718 101
REPEAT 719 101
REPEAT 720 101
REPEAT 721 101
REPEAT 722 101
REPEAT 723 101
REPEAT 724 101
REPEAT 725 101
REPEAT 726 101
REPEAT 727 101
REPEAT 728 101
REPEAT 729 101
REPEAT 730 101
REPEAT 731 101
REPEAT 732 101
REPEAT 733 101
REPEAT 734 101
REPEAT 735 101
REPEAT 736 101
REPEAT 737 101
REPEAT 738 101
REPEAT 739 101
REPEAT 740 101
REPEAT 741 101
REPEAT 742 101
REPEAT 743 101
REPEAT 744 101
REPEAT 745 101
REPEAT 746 101
REPEAT 747 101
REPEAT 748 101
REPEAT 749 101
REPEAT 750 101
REPEAT 751 101
REPEAT 752 101
REPEAT 753 101
REPEAT 754 101
REPEAT 755 101
REPEAT 756 101
REPEAT 757 101
REPEAT 758 101
REPEAT 759 101
REPEAT 760 101
REPEAT 761 101
REPEAT 762 101
REPEAT 763 101
REPEAT 764 101
REPEAT 765 101
REPEAT 766 101
REPEAT 767 101
REPEAT 768 101
REPEAT 769 101
REPEAT 770 101
REPEAT 771 101
REPEAT 772 101
REPEAT 773 101
REPEAT 774 101
REPEAT 775 101
REPEAT 776 101
REPEAT 777 101
REPEAT 778 101
REPEAT 779 101
REPEAT 780 101
REPEAT 781 101
REPEAT 782 101
REPEAT 783 101
REPEAT 784 101
REPEAT 785 101
REPEAT 786 101
REPEAT 787 101
REPEAT 788 101
REPEAT 789 101
REPEAT 790 101
REPEAT 791 101
REPEAT 792 101
REPEAT 793 101
REPEAT 794 101
REPEAT 795 101
REPEAT 796 101
REPEAT 797 101
REPEAT 798 101
REPEAT 799 101
REPEAT 800 101
REPEAT 801 101
REPEAT 802 101
REPEAT 803 101
REPEAT 804 101
REPEAT 805 101
REPEAT 806 101
REPEAT 807 101
REPEAT 808 101
REPEAT 809 101
REPEAT 810 101
REPEAT 811 101
REPEAT 812 101
REPEAT 813 101
REPEAT 814 101
REPEAT 815 101
REPEAT 816 101
REPEAT 817 101
REPEAT 818 101
REPEAT 819 101
REPEAT 820 101
REPEAT 821 101
REPEAT 822 101
REPEAT 823 101
REPEAT 824 101
REPEAT 825 101
REPEAT 826 101
REPEAT 827 101
REPEAT 828 101
REPEAT 829 101
REPEAT 830 101
REPEAT 831 101
REPEAT 832 101
REPEAT 833 101
REPEAT 834 101
REPEAT 835 101
REPEAT 836 101
REPEAT 837 101
REPEAT 838 101
REPEAT 839 101
REPEAT 840 101
REPEAT 841 101
REPEAT 842 101
REPEAT 843 101
REPEAT 844 101
REPEAT 845 101
REPEAT 846 101
REPEAT 847 101
REPEAT 848 101
REPEAT 849 101
REPEAT 850 101
REPEAT 851 101
REPEAT 852 101
REPEAT 853 101
REPEAT 854 101
REPEAT 855 101
REPEAT 856 101
REPEAT 857 101
REPEAT 858 101
REPEAT 859 101
REPEAT 860 101
REPEAT 861 101
REPEAT 862 101
REPEAT 863 101
REPEAT 864 101
REPEAT 865 101
REPEAT 866 101
REPEAT 867 101
REPEAT 868 101
REPEAT 869 101
REPEAT 870 101
REPEAT 871 101
REPEAT 872 101
REPEAT 873 101
REPEAT 874 101
REPEAT 875 101
REPEAT 876 101
REPEAT 877 101
REPEAT 878 101
REPEAT 879 101
REPEAT 880 101
REPEAT 881 101
REPEAT 882 101
REPEAT 883 101
REPEAT 884 101
REPEAT 885 101
REPEAT 886 101
REPEAT 887 101
REPEAT 888 101
REPEAT 889 101
REPEAT 890 101
REPEAT 891 101
REPEAT 892 101
REPEAT 893 101
REPEAT 894 101
REPEAT 895 101
REPEAT 896 101
REPEAT 897 101
REPEAT 898 101
REPEAT 899 101
REPEAT 900 101
REPEAT 901 101
REPEAT 902 101
REPEAT 903 101
REPEAT 904 101
REPEAT 905 101
REPEAT 906 101
REPEAT 907 101
REPEAT 908 101
REPEAT 909 101
REPEAT 910 101
REPEAT 911 101
REPEAT 912 101
REPEAT 913 101
REPEAT 914 101
REPEAT 915 101
REPEAT 916 101
REPEAT 917 101
REPEAT 918 101
REPEAT 919 101
REPEAT 920 101
REPEAT 921 101
REPEAT 922 101
REPEAT 923 101
REPEAT 924 101
REPEAT 925 101
REPEAT 926 101
REPEAT 927 101
REPEAT 928 101
REPEAT 929 101
REPEAT 930 101
REPEAT 931 101
REPEAT 932 101
REPEAT 933 101
REPEAT 934 101
REPEAT 935 101
REPEAT 936 101
REPEAT 937 101
REPEAT 938 101
REPEAT 939 101
REPEAT 940 101
REPEAT 941 101
REPEAT 942 101
REPEAT 943 101
REPEAT 944 101
REPEAT 945 101
REPEAT 946 101
REPEAT 947 101
REPEAT 948 101
REPEAT 949 101
REPEAT 950 101
REPEAT 951 101
REPEAT 952 101
REPEAT 953 101
REPEAT 954 101
REPEAT 955 101
REPEAT 956 101
REPEAT 957 101
REPEAT 958 101
REPEAT 959 101
REPEAT 960 101
REPEAT 961 101
REPEAT 962 101
REPEAT 963 101
REPEAT 964 101
REPEAT 965 101
REPEAT 966 101
REPEAT 967 101
REPEAT 968 101
REPEAT 969 101
REPEAT 970 101
REPEAT 971 101
REPEAT 972 101
REPEAT 973 101
REPEAT 974 101
REPEAT 975 101
REPEAT 976 101
REPEAT 977 101
REPEAT 978 101
REPEAT 979 101
REPEAT 980 101
REPEAT 981 101
REPEAT 982 101
REPEAT 983 101
REPEAT 984 101
REPEAT 985 101
REPEAT 986 101
REPEAT 987 101
REPEAT 988 101
REPEAT 989 101
REPEAT 990 101
REPEAT 991 101
REPEAT 992 101
REPEAT 993 101
REPEAT 994 101
REPEAT 995 101
REPEAT 996 101
REPEAT 997 101
REPEAT 998 101
REPEAT 999 101
REPEAT 1000 101
REPEAT 1001 101
REPEAT 1002 101
REPEAT 1003 101
REPEAT 1004 101
REPEAT 1005 101
REPEAT 1006 101
REPEAT 1007 101
REPEAT 1008 101
REPEAT 1009 101
REPEAT 1010 101
REPEAT 1011 101
REPEAT 1012 
```


